

Independence of forensic science

The third appeal of six people against life sentences imposed 15 years ago raises sombre questions about the management of forensic services, not to mention the quality of British justice.

THE application of scientific techniques to the detection of crime and the conviction of criminals has a long and even a spectacular history, with the result that the achievements of forensic science have spawned a whole genre of airport fiction. But real life, usually less glamorous, is full of pitfalls for the unwary. So much is plain from the third re-hearing now under way by the British Appeal Court of the evidence for the conviction in 1975 of six Irishmen for having killed 24 people (and injured more than 120) by exploding a bomb in a crowded drinking-house in Birmingham. The 'Birmingham six', as they are known, were convicted and imprisoned for life on evidence of two kinds: purported confessions to the police who had arrested them and evidence (accepted by the trial court as valid) that two of them had been found on forensic investigation to have traces of explosives on their right hands.

What now emerges is that the forensic evidence presented at the original trial was flawed, and has for practical purposes been withdrawn. A chemical procedure for identifying nitroglycerine has now been found to have been inadequately controlled: materials other than nitroglycerine (including soap) can yield positive results. Results of an investigation by gas chromatography, potentially a more sensitive technique, were acknowledged in evidence last week to be similarly capable of confusion by other materials (and, even so, were positive for the *left* hand of only one of the two supposedly contaminated defendants). But forensic science has salvaged some of its reputation by producing evidence (still formally to be accepted by the Appeal Court) that the disputed confessions produced by the police consisted of writings (by the interrogating officers) that were neither contemporaneous nor consecutive.

Two important issues stem from these alarming disclosures, of which the chief is more than a mere embarrassment to the British government and judicial system. If, as seems probable, the Birmingham six are now released after 16 years in jail, this will be yet another case in which zealous and over-zealous policemen have been prepared to concoct evidence in the pursuit of a conviction. That many of these cases have involved supposed terrorism by the Irish Republican Army has not, to say the least of it, simplified the management of Ulster's problems.

The forensic issue is, or should be, more tractable. In Britain, the Forensic Science Service has traditionally been run by the Home Office as a common service for Britain's police forces. But from 1 April, it will become an executive

agency — expected to pay its way but not to make a profit. The forensic service, which has six laboratories south of the Scottish border, will take business from all potential customers: police forces, but also defendants before the courts. The danger is that the service will be yet more dependent on the wishes of its customers, including policemen out for a conviction.

That is why a still more radical change is required. It was claimed (by counsel for the appellants) at last week's hearing that evidence that the original test for nitroglycerine was not reproducible had been withheld from the trial court — and from defence lawyers. It is only fair (to the forensic service 15 years ago) to acknowledge that the court might have learned the truth if the defence lawyers had asked the right questions (and only fair to them to acknowledge that lawyers are rarely trained in the niceties of laboratory technique).

Yet there can be few occasions when the results of forensic investigations are as clear-cut as the airport literature would have the rest of us believe, let alone when lawyers appreciate what the uncertainties of forensic investigations mean. So much has been well illustrated by the arguments over the unreflective use of DNA fingerprinting by some US courts (see *Nature* 339, 501; 1989). The only durable solution is that forensic services everywhere should become the servants of the courts at which their evidence is presented, not of those who commission their investigations.

In British circumstances, the simple solution is that they should be a child of the Crown Prosecution Service, which has (among other things) responsibility for investigating the quality of evidence against accused persons. That is the solution advocated by Lord Scarman, Britain's most distinguished ex-judge. But why not go further, and make them dependent on the courts as such, as are the bailiffs who go about collecting debts? The snag, for the government, would be that the prospect that the forensic services would be self-financing might then melt away, but when the reputation of the judicial system has been undermined by faulty evidence, that hardly seems a substantial consideration. □

Technology untied?

Plans to liberalize British telecommunications would be more commendable if they were more radical.

SINCE soon after the time of Alexander Graham Bell, govern-



ments have asserted a right to regulate the uses made of his invention and its successors, and for two kinds of reasons. First, as in the pre-*glasnost* Soviet Union, telephones provide people with potentially subversive opportunities for communication, which governments seek to prevent either by not providing a workable service or by listening to the conversations that telephones generate. But even elsewhere, telecommunications services are not free. Because telephone networks are natural physical monopolies, public regulation of even private enterprises is unavoidable. It is not merely that unregulated communications networks might exploit their obligatory customers financially, but that they might discriminate among them on political or other such grounds. During last year's elections in Eastern Europe, for example, opposition parties complained loudly that the state providers of communications gave them too little access to the inadequate facilities that there were.

Given all that, the British government deserves a little credit for having decided, last week, that there should now be a further liberalization of the regulatory framework of British telecommunications. (During the same week, it also decided that all airlines can in future apply for the right to land at London's principal airport, which is a further sign of liberal inclinations — see *Nature* 349, 638; 21 February 1991). Yet what the government now proposes is the least of what it should have done a decade ago, when 51 per cent of the publicly owned telecommunications monopoly (called British Telecom) was sold to the public. And the snail's pace of regulatory change promises to lag further behind the pace of technical change — which, ironically, the government seeks further to accelerate.

In the new regime, as in the old, operators of telecommunications services will have to be licensed, which is proper, while the scope of the services they provide and the prices at which they sell them will continue to be overseen by a regulatory agency, called Oftel. Ten years ago, the government licensed British Telecom to provide a comprehensive telecommunications service (nationally as well as overseas) and encouraged a second (called Mercury) to build a competing network. But the government was also then grappling with (even cheerleading) the prospect that the opportunities in cable television would quickly thread the country with an independent broad-band telecommunications network. (The prospect might have become reality had the government not fortuitously moved the financial goalposts.) Meanwhile, all concerned have evidently been astonished by the pace of development of mobile networks based on UHF radio.

What is now proposed is a modestly sensible loosening of this framework. Newcomers to the communications business will in principle be licensed (but not immediately). Meanwhile, operators of mobile communications systems and of cable networks will be allowed to use their facilities to provide fixed-point, local and even trunk communications. Private organizations wishing to set up communications systems by means of Earth satellites will be allowed to do so, provided that their channels are not connected to a fixed network at either end. Both British Telecom and Mercury will have to allow others access to their networks on terms to be

negotiated, and will also be allowed to continue subsidizing their domestic services and capital investments by making uneconomically high international charges (which are nevertheless to be reduced by 10 per cent). At least for seven years, perhaps for ten, British Telecom will not be allowed to use its network for distributing television signals.

The last of these restrictions points directly to the technical anomaly now being created. With broad-band communications just around the corner, it would make technical sense that there should be a single broad-band network based on optical-fibre cables — and one set of holes in urban streets. But would not that be a recipe for domination of the market by its most likely provider, British Telecom? Not if the regulatory framework required that provider to function as a common carrier, charging transparent prices for access to the network and having itself no commercial interest in the uses made of the network — voice communications, data transmission or even entertainment. The British government's difficulty is that, even if it is not influenced by the 49 per cent of the company it has retained, it has a kind of moral commitment to those who own the rest of the stock to ensure that the company is not emasculated commercially. But the perpetually managed profitability of British Telecom is hardly a more laudable objective than those that embarrassed successive governments when British Telecom was a nationalized industry. □

Women without men

A row about the insemination of unmarried women scorns human nature and undervalues human ingenuity.

WITH the Gulf War over for the time being, the British tabloid press and even some members of the House of Commons (MPs) are making a fuss about the revelation that some women (there are at least two) have persuaded a fertility clinic to enable them to become pregnant with sperm from a donor-bank. "Virgin birth" features in most headlines, but there is also a drum-beat of thundering about the sanctity of the nuclear family, and yet another outburst of anxiety about the new human embryology (misplaced because artificial insemination by donor, or AID, is old hat).

What seems to have been overlooked is that many unmarried women have traditionally sought to bear children, and to bring them up on their own, without the assistance of fertility banks. Nobody could possibly tell what proportion of the large number of unmarried mothers in Britain are in this category, but it would not be negligible. Two parents are better than one, and not simply for the obvious prudential reasons (one might die). But these days it takes a brave newspaper or MP to say that a woman without a mate must not become pregnant. And who is to say that mating with a passing acquaintance is preferable to a fertility bank? Is it possible that the indignation so suddenly aroused stems from the British government's plan to insist that absent fathers make maintenance payments on behalf of their children? Will the fertility banks be responsible? □

Human Frontier effort gets promise from NIH

- Institutes offer in-kind contributions
- Psychological and political benefits

Tokyo & Washington

A LETTER from a senior official at the US National Institutes of Health (NIH) indicates that NIH have decided to contribute to the Human Frontier Science Program (HFSP), the ground-breaking Japanese effort to support international research on the brain and molecular biology through a foundation based in Strasbourg, France.

Although the US contribution will be relatively small and will be confined to support of researchers working in the United States, the NIH decision comes at a very opportune time for the Human Frontier programme — it is nearing the end of a three-year trial period in which nearly all its financial support has been provided by Japan, and it is now entering a critical phase of reassessment. Furthermore, the promise of an NIH contribution marks a remarkable reversal in the US attitude towards the programme, which was greeted with suspicion by the United States when it was first announced by Japan's Ministry of International Trade and Industry (MITI) in 1986.

The NIH decision was announced by Edward Rall, its deputy director of intramural research, in a letter dated 26 February to Sir James Gowans, the British secretary-general of the programme in Strasbourg. On Monday, however, there seemed to be some confusion in NIH as to just how far the decision-making process had gone. Jerome Green, who is NIH's representative on the Frontier programme's board of trustees, said there had been no final decision on making contributions in kind, "letters and rumours to the contrary".

Rall was out of town and not available for comment, but his letter clearly indicates that the NIH policy is set. The letter states that NIH will make an 'in kind' contribution for HFSP fellowships as follows: "Any fellow who is duly selected by the Fellowship Panel and approved by the Council of Scientists, and who plans to work at any laboratory in the National Institutes of Health, or the Alcohol, Drug Abuse and Mental Health Administration, will be supported by the National Institutes of Health in accordance with the same stipend level as the HFSP would provide." A notation at the bottom of the letter indicates that copies were sent to several science policy officials in Washington, although not to Green.

Such in-kind funding from the NIH will probably not cost the agency a great deal of money. In fiscal year 1990, only a handful of scientists were awarded HFSP fellowships to work at NIH laboratories, and with a similar number of awards expected this year, the size

of the US contribution will amount to only a few hundred thousand dollars at most.

That is just a small percentage of the total cost. In fiscal year 1990, the Frontier programme awarded 29 grants of up to \$400,000 a year to international teams, 80 postdoctoral fellowships worth \$40,000 a year and ten workshop grants of \$80,000 each. Nearly \$20 million, or more than 75 per cent of the \$26 million budget for the programme in fiscal year 1990, was provided by MITI and Japan's Science and Technology Agency (STA). The remainder was made up from \$4.2 million carried over from fiscal year 1989, and relatively small contributions from four of the other six summit nations that belong to the HFSP: France (\$1.24 million), Italy (\$0.24 million), Canada (\$0.20 million) and Germany (\$0.12 million). (The United Kingdom and the United States have made no cash contributions.)

But despite the relatively small size of the forthcoming NIH contribution, psychologically and politically it is very important.

Japanese backers of HFSP have been seriously concerned that the programme might run into funding difficulties in Japan if other summit nations, particularly the United States, did not contribute more. At the end of fiscal year 1991 (31 March 1992), the programme comes to the end of its three-year trial period, and the Japanese ministry of finance must decide if — and on what scale — to continue. Professor Akiyoshi Wada, one of the two Japanese members of the HFSP council of scientists, said last week

that the lack of contributions from other countries was threatening the future of the programme.

The failure of the programme would be a great loss not only for scientific reasons, he said. Many science and government officials in the summit nations think highly of the programme. For instance, Philip Schambra, director of the Fogarty International Center at NIH, says that HFSP fulfils an important function in encouraging international co-operation in biological research and should be expanded. But Wada points out that the programme has meant even more than that to Japan. Although few people in the West realize it, he said, the programme has had tremendous beneficial effects in Japan by exposing some of the domestic bureaucratic barriers facing the internationalization of science. He likens Frontiers to the coming of the US 'black ships' led by Commodore Matthew C. Perry in the last century, which opened up Japan to the West after over two centuries of isolation.

So Wada was relieved to hear the news that NIH and Switzerland (see below) will contribute to HFSP, and says he is now more optimistic about the future.

In January, the HFSP board of trustees began the process of reassessing the programme. That process will be concluded in June — just in time for MITI and STA to make budget applications for fiscal year 1992. The board is expected to recommend a further two-year trial period rather than giving its full-blooded support because it is too early to assess the scientific achievements of the programme. Wada and Japanese government officials in charge of the programme hope that the votes of confidence from NIH and the Swiss will be enough to convince the Ministry of Finance not only that HFSP is worth continuing but also that it deserves increased support.

David Swinbanks &
Robert Pool

Swiss sign up also

SWITZERLAND has been officially accepted into Japan's Human Frontier Science Program (HFSP), becoming the first nation to join the programme that was not one of the seven Summit nations that were the original members.

The newest member will contribute \$400,000 a year, making it the third largest contributor after Japan and France. The acceptance of Switzerland into HFSP may stimulate other non-Summit nations to sign up for the programme, which supports international research on the brain and molecular biology.

Switzerland, Sweden and Australia have long expressed an interest in joining the Frontiers programme, but an initial offer from these countries last year was rebuffed by the HFSP board of trustees. The deal fell through partly because the

board considered the amount offered (about \$150,000 a year) to be too small. A second problem was that the Japanese officials negotiating the deal had offered preferred treatment to nations willing to make contributions, says Katherine Bick, who represented the United States on the board of trustees at that time. She says the board told the Japanese government that the grants would have to be made strictly on merit. And so the matter went back to the negotiating table (see *Nature* 345, 9; 1990).

Japanese government officials say they have not heard from Sweden or Australia recently and the "ball is now in their court", but Switzerland is now a fully fledged member of the programme. Two Swiss have been nominated to the board of trustees, and a Swiss scientist will join the council of scientists when it meets in Strasbourg next week.

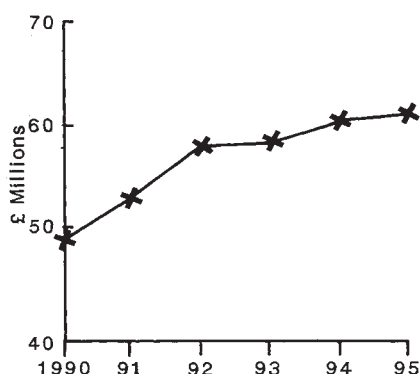
D.S. & R.P.

Does Europe need the LHC?

London

SIR David Phillips, chairman of the Advisory Board for the Research Councils (ABRC), last week angered British physicists by questioning in public the need for a separate European programme in particle physics. His comments brought out into the open a long-running debate about whether the world needs two science projects costing thousands of millions of dollars each to address essentially the same question.

Speaking to the House of Lords Science and Technology Committee, Phillips said



Phillips's headache — the planned UK subscription to CERN, 1990–95. Source: SERC.

that particle physics should be coordinated at a global level, but the “competitive spirit” of rival groups of physicists has won through. Science advisers in France and Germany, as well as in Britain, are concerned about the level of planned expenditure at CERN, the European particle physics laboratory in Geneva, Phillips said. The SFr 1730 million (about £680 million) cost of building a new accelerator, the Large Hadron Collider (LHC), is particularly worrying, he said, because the plan coincides with a similar US initiative — the \$8,300-million Superconducting Super Collider (SSC) project.

The British subscription to CERN, at more than £50 million a year, is the largest single component of the UK science budget. So, as he is chairman of the ABRC, which advises the Education and Science Secretary on the size and distribution of the science budget, Phillips' concern about future increases in the CERN subscription is understandable. But his comments come at a sensitive time: the CERN member nations must next year decide whether or not to go ahead with the LHC. Carlo Rubbia, director-general of CERN, last week declined to comment on the criticisms, preferring to let British physicists rally to CERN's cause.

CERN's supporters argue that if there is an undesirable overlap between the LHC and SSC projects, the fault lies with the United States. Sir William Mitchell, former chairman of the UK Science and Engineering Research Council (SERC), says that CERN and the British government were

pushing hard for a world strategy in particle physics in the late 1980s. But “it takes two to tango”, he says — the United States was determined that the next stage in its particle physics programme would be a US-dominated project. Sandy Donnachie, chairman of SERC's Nuclear Physics Board, which handles the UK CERN subscription, says the US attitude was: “we must regain the initiative from the Europeans”. Nevertheless, Roy Schwitters, director of the SSC laboratory, responds that European physicists will be involved in experiments on the SSC, and are represented in his laboratory's peer review committees.

Mitchell, who is now president of CERN's governing council, believes the LHC project will not increase CERN's budget dramatically, as the new collider can be built around CERN's current centrepiece, the Large Electron-Positron collider, using some of the same equipment. Donnachie agrees: “if Europe were to buy into the SSC, it would finish up costing us more than the LHC.”

Chris Llewellyn Smith, from the University of Oxford, says that competition between rival groups can be a spur to better science. He expects the LHC to be completed before the SSC, but says the LHC's greater versatility is in any case a powerful argument against simply pinning European hopes on a contribution to the SSC.

The primary function of both the SSC and

OIL-WELL FIRES

Smog alert in Kuwait City

London

By concentrating on the large-scale environmental impact of the hundreds of oil-well fires now blazing in Kuwait (see *Nature* 350, 11; 7 March 1991), scientists and the media may have neglected a more immediate problem — the threat to public health in Kuwait posed by toxic fumes released into the atmosphere.

At present, plumes of smoke and fumes are rising rapidly into the atmosphere, so dangerous concentrations of toxic chemicals are not building up at ground level in populated areas. But David Shillito, from the consultant chemical engineers Cremer and Warner, fears that Kuwait City will face a serious air pollution problem in the coming months, as seasonal winds begin to blow. When the wind speed hovers around 12 metres per second, Shillito says, the smoke plumes will be “blown over”, but will not disperse rapidly.

In June and early July, a prevailing wind known locally as the Shamal blows over the northern Kuwaiti oil fields towards the capital. Shillito, who worked in the Persian Gulf for many years, says that the air quality in Kuwait City may then be as bad as the infamous London smogs of the 1950s, for periods of 2–3 days at a time. Oil-fire smoke

the LHC is to collide two beams of protons. Proton collisions may yield the much sought after Higgs boson, predicted by current ideas in theoretical physics. The SSC will operate at a higher energy, but Llewellyn Smith says that greater intensity of the LHC proton beams should compensate for this, in part.

The LHC will also be able to operate as an electron-proton collider, providing the next step beyond the German HERA machine, where experiments start later this year. A third role is also planned, using the LHC to collide heavy ions. By forcing heavy nuclei



together, physicists may be able to duplicate the ‘soup of quarks’ that must have existed early in the evolution of the Universe, Llewellyn Smith says. A separate \$300-million US machine is planned to conduct similar experiments (but at a fraction of the energy the LHC will deliver) at the Brookhaven National Laboratory in New York State.

Peter Aldhous

contains a cocktail of noxious chemicals, from sulphur dioxide and other simple respiratory irritants, to carcinogens such as benzene and polyaromatic hydrocarbons.

Keith Browning, director of research at the UK Meteorological Office, says that his group has not been asked to extend its modelling of the environmental effects of the fires to assess in detail the likelihood of local air pollution events, but says that this “must represent a threat”. Browning's group has considered the formation of ground-level ozone, by the action of sunlight on nitrogen oxides. Near the source of each fire, little light will penetrate the black smoke, but Browning says intermittent photochemical smogs are possible where the plumes fall back to the ground.

The Kuwaiti health ministry has asked the Geneva-based World Health Organization (WHO) for help in the aftermath of the war, and the air pollution threat is a top priority. WHO expects to announce this week that it is sending a team of clinical experts to Kuwait.

Another WHO concern is the serious risk of disease epidemics in Iraq, where allied bombing has caused severe disruption to water supplies and sanitation.

Peter Aldhous

Amgen scores a knockout

Washington

DESPITE expectations to the contrary, Amgen Inc. has emerged as the outright winner in its bitter dispute with Genetics Institute over US patent rights to erythropoietin (EPO). The US Court of Appeals for the Federal Circuit ruled that key claims in Amgen's patent were valid and enforceable and declared Genetics Institute's patent invalid.

The decision guarantees a market monopoly for Amgen. As a result, Genetics Institute will be unable to sell its version of EPO in the United States through its licensee Chugai Pharmaceutical Co. Ltd. EPO, which stimulates the growth of red blood cells, is proving to be a valuable therapy for patients on kidney dialysis, who often develop anaemia. AIDS patients with anaemia may also benefit from EPO.

News of the court decision took Wall Street by surprise and sent the company's shares up 12 per cent to close at \$113. Just three weeks before the court's EPO decision, Amgen won US Food and Drug Administration (FDA) approval to market granulocyte-colony stimulating factor (G-CSF), another biotechnology drug. The colony stimulating factor enhances white blood cell production and is expected to find wide use among cancer patients and others who are suffering from immune deficiency.

Gordon Binder, Amgen's chairman and chief executive officer, stated that the court decision in the EPO case "vindicates our position completely". He believes that the ruling will encourage original research and development by reassuring the pioneers within the biotechnology industry that they can expect adequate patent protection for their inventions. Gabriel Schmergel, president and chief executive officer of Genetics Institute said "we are very disappointed with the decision" but would not indicate whether the company will take any further action.

The protracted patent dispute spanned three-and-a-half years and concerns conflicting US patent claims to EPO, a natural protein that is made in the kidney. In June 1989, the FDA approved Epogen, Amgen's genetically engineered version of EPO, for the treatment of anaemia in patients with chronic kidney failure. Estimated US sales of Epogen for 1992 are expected to top \$400 million. Along with marketing approval, FDA awarded Epogen "orphan drug" status under an act that provides companies with incentives, including a guaranteed seven-year market monopoly, to develop drugs to treat rare diseases.

At the time, both Amgen and Genetics Institute held EPO patents. Amgen won its patent for discovery of the gene that produces EPO, whereas the Genetics Institute patent was awarded for isolation of the EPO molecule from human urine, as well as for preparation of a recombinant version

of the protein. A court in Massachusetts, which heard the patent case in 1989, validated the central claims in both Amgen's and Genetics Institute's patents but declared that they were mutually infringing (see *Nature* 342, 846; 1989). Observers expected that the appeals court in Washington, DC, would uphold the Massachusetts ruling and that the continued stalemate would force the companies to cross-license.

Although the federal appeals court affirmed Amgen's patent as valid and enforceable, Genetics Institute was dealt a severe and unexpected blow when the court struck down its patent claims to EPO. The court ruled that Genetics Institute and its licensee, Chugai, had failed to show that they had purified EPO, whether recombinant or urinary-derived, with the claimed specific activity using the methods disclosed in the patent. Genetics Institute had argued that its

claim to EPO was stronger because Edward F. Fritsch of Genetics Institute first conceived the strategy to identify and isolate the EPO gene back in 1981, even though Fu-Kuen Lin of Amgen was first to accomplish this goal in September 1983, followed by Fritsch eight months later. The appeals court determined, however, that Fritsch "had no more of a conception in 1983 than he did in 1981 because he did not know the sequence of the gene encoding EPO."

Peter Drake, stock analyst with Vector Securities, paid tribute to Amgen's management team, who he said had the long-term vision, energy and commitment to defend its patent position "in the face of significant pressure from Wall Street and other circles to cross-license and put this [the patent dispute] behind them". Amgen is now in the enviable position of having "two of the crown jewels in the biotechnology industry", says Drake, who estimates that the company will have \$1,000 million in sales by 1993.

Diane Gershon

NATURAL HISTORY MUSEUM

Innovative ecology exhibits

Leafy environment — shrinking down to 8,000 times smaller, visitors to the ecology exhibition can take a tour around the 'inside' of a leaf.

A TOWERING glass greenhouse inside the Natural History Museum in London sets the scene for an impressive yet controversial permanent exhibition on ecology that opened on 7 March. It is the first of five new exhibitions scheduled to open over the next few years and sponsored by the museum's independent Development Trust Appeal. The exhibition marks an important shift away from government funding, but the in-house team largely responsible for the ex-

hibition is to be disbanded as a consequence of the museum's cost-cutting 1990–91 corporate plan. The visually exciting exhibition, which was a result of careful market research, has not met with universal praise, however. Its interactive video and spectacular effects have largely replaced traditional exhibits, something that a curator at another UK museum described as "appalling".

Henry Gee, London

Back to the drawing board

Washington

A two-year-old experiment in the investigation of scientific misconduct has blown up in the face of the US National Institutes of Health (NIH). The special office that NIH set up to investigate allegations of scientific misconduct has come under courtroom and congressional attack, and late last year a district judge in Wisconsin ruled that the office's procedures were illegal.

Now, as a host of similar lawsuits raise the prospect of NIH's misconduct rules being declared illegal in the rest of the country, the agency has decided to start the process again from the beginning. Officials say that within weeks NIH will publish in the *Federal Register* proposed new rules and procedures for the Office of Scientific Integrity (OSI), the department in question.

The move comes some three months after a crucial decision in the case of James Abbs, a researcher at the University of Wisconsin, who had been under investigation for alleged data fabrication (see *Nature* 349, 95; 10 January 1991). The judge in that case ruled that NIH's investigation was illegal because OSI's internal procedures had been created *in camera*. The law demands that such procedures be published for public comment before they become law. At least two other lawsuits citing the Abbs decision have been filed since then (see *Nature* 349, 357; 31 January 1991).

NIH intend to concede the issue and comply with the Administrative Procedures Act, says William Raub, acting director of NIH. The act requires that any proposed rules that could injure people outside government must be published with a comment period during which members of the public can suggest changes. Because OSI has taken the authority to go beyond simple investigations to the point of imposing sanctions (such as stopping a grant), it is subject to the act.

During the forthcoming rule-making process, which typically lasts for one to two years, OSI will have no legal procedures under which to operate. Nevertheless, NIH intend to continue investigating misconduct cases, but on a 'case-by-case basis' as they did before OSI was established in 1989.

"Before OSI was started, we thought we could handle investigations without written procedures", Raub says. "We thought the cases would be so few and so different that we wouldn't have to write things down. We were wrong."

OSI was created to bring order to the investigation process. But critics soon noted that the OSI, like NIH before it, tended to make up procedures as it went. (In the case of AIDS researcher Robert Gallo, for example, OSI asked the National Academy of Sciences to help it to set up a special advisory panel.) Faced with bitter complaints of arbitrary and capricious rule-making, OSI decided to formalize and publish its current

rules (see *Nature* 346, 9; 1990).

That, however, turned out to be a fatal mistake. A federal office with no rules may draw criticism, but an office that creates rules that conflict with the Administrative Procedures Act risks courtroom challenge, as NIH discovered.

Abandoning its existing procedures may lift the legal cloud over OSI for the moment, but it is unlikely to end its problems. If the office continues to impose sanctions, further lawsuits can be expected. Worse still, there is the possibility that researchers whose cases have already been closed may sue on the grounds that they were investigated under illegal rules. But Raub says that he can see no better way out. "Doing nothing is not an option", he says.

Robert Charrow, a Washington attorney, suggests a solution. If NIH would give up sanctions, he says, their legal problems could be over. "They ought to announce that all they will do is investigate, and then simply use the accepted hearing processes that are already in the books." NIH may eventually have to adopt that course. But until it is forced to abandon sanctions, the agency is eager to retain the right to halt research funding when the evidence indicates misconduct, if only for the sake of appearances.

Although the final decision on when or

CONFLICT OF INTEREST

Black eye for NIH

Washington

OFFICIALS at the National Institutes of Health (NIH) bungled a case of apparent conflict of interest, even after Congress pointed it out to them almost a year ago, a new investigation has revealed. The case itself is relatively pedestrian — Prem Sarin, deputy chief of AIDS pioneer Robert Gallo's NIH laboratory, is alleged to have illegally consulted with industry and set up a secret bank account to hide payments. But NIH's handling of the allegations raises serious questions about the agency's ability to police its own researchers.

At a hearing last week, Representative John Dingell (Democrat, Michigan), the firebrand chairman of the Energy and Commerce Committee's oversight and investigations subcommittee, asked NIH acting director William Raub, Gallo and two other NIH officials to explain their actions in the case.

Dingell noted that possible conflict of interest in Sarin's activities had first been brought to NIH's attention in April last year at a hearing on misuse of funds by Zaki Salahuddin, another researcher in Gallo's lab (see *Nature* 345, 99; 1990). Salahuddin later pleaded guilty and was sentenced in district court (see *Nature* 348, 8; 1990).

At the May hearing, Dingell revealed that a Food and Drug Administration (FDA)

whether to publish new OSI procedures is now in the hands of the Justice Department, where lawyers are still considering their options, NIH officials are already writing the proposed rules. Raub says that the first draft will essentially codify the existing procedures. "We'll use that as a starting point", he says.

Concern in the scientific community is likely to centre on the issue of 'due process' — the protections and civil liberties guaranteed in US courts. Although NIH officials point out that OSI is not a court, several lawsuits have argued that an OSI investigation is as arduous as a criminal investigation, and can take far longer (in some cases more than two years). In particular, researchers have complained that the OSI rules do not give them the right to confront their accuser face-to-face and to see the evidence against them. But NIH, in an effort to protect whistleblowers, intend to fight that point.

Once NIH publish the new OSI procedures, there will be a comment period (typically 90 days) during which the agency will accept any public input. OSI will then re-write the rules in line with public consensus and legal concerns. Finally, a finished version of the rules will be published in the *Federal Register* and become law, ready for the next batch of misconduct cases. As for cases now under way or already closed, the courts appear to be the next battleground.

Christopher Anderson

employee had written to the subcommittee to complain that Sarin had represented Responsif, a pharmaceutical company, at a 1985 FDA meeting, despite the fact that he was an NIH employee. Although Sarin, who was present at the Dingell hearing because he was Salahuddin's immediate supervisor, testified that his actions had been legal under NIH rules at the time, other NIH officials at the hearing were unable to confirm that and said that they would look into the matter.

What the officials actually did, an investigation by the congressional General Accounting Office (GAO) has revealed, was to ask Gallo to talk to Sarin about the matter. Sarin assured Gallo that "everything was taken care of", Gallo told the investigators. Dingell noted that this was exactly the way Salahuddin had successfully deflected suspicion when Gallo had questioned him on his own allegations of wrongdoing a year previously. "Apparently, these are magic words", Dingell wryly commented.

"It is discouraging", Dingell said, "that NIH did nothing to investigate the Sarin matter between the time of the April subcommittee hearing and December 1990", when his staff requested certain NIH documents that clearly showed irregularities in Sarin's consulting. Only then, when NIH were virtually handed evidence of wrongdoing

EUROPEAN RESEARCH

Commission bids for power

ing, did officials investigate the case. NIH removed Sarin from his position earlier this year (see *Nature* 349, 95; 1991) and the case is now being investigated by the agency's Inspector General and the US Attorney's office.

"It is not the business of the subcommittee to supervise the daily workings of the NIH. But if we have to do it, we will", Dingell threatened.

The most damaging evidence in the Sarin case revolves around an NIH fund — known as the Foundation for the Advanced Education in the Sciences (FAES) — that was used to support visiting scientists in Gallo's laboratory. According to the GAO investigators, Sarin set up a private bank account at a local bank with a similar name — FAES/NERIC. These initials, however, stood for the "Family Account for the Education of the Sarin Children/Neil and Eric". Sarin instructed ASTA Pharma, a German pharmaceutical company, to make out a consulting cheque for \$25,000 to FAES (an illegal payment, GAO said), which he deposited in the private account, along with other consulting payments. Sarin's lawyer, W. Neil Eggleston, says that Sarin has violated no laws, but he declines to explain the details of the questioned bank account.

NIH did not spot the violations, Raub testified, because officials had not thoroughly examined Sarin's files and were not sure that he had been given correct advice when he sought approval for the consulting. "In hindsight, the level of trust was too high", he said. "One of the things we need to develop across NIH is a greater talent for selective suspicion."

Gallo testified that, although he had been asked to question Sarin about the allegations, he could do little more than accept Sarin's denial.

"If the subcommittee and the NIH know about the allegations of misconduct, I don't know what I'm doing investigating it", Gallo argued. "I'm a scientist, not a policeman."

It is not clear exactly what NIH have done to prevent and investigate conflict of interest abuses in the future. "Do you still go by the same seat-of-the-pants review process that brought about this sorry situation?", Dingell asked Richard Adamson, Gallo's superior at the National Cancer Institute (NCI).

"No", Adamson responded. He was, however, unable to specify any significant new protection in the NIH procedures prompted by the Sarin case. Raub testified that, since the Salahuddin hearing, NIH have improved its review of applications for outside consulting and started an NIH-wide inventory of equipment. But most effort seems to have been spent in ensuring that existing procedures are being followed. NCI employees have been given seminars on existing conflict-of-interest rules, and outside companies will be sent copies of the relevant regulations, so that they can watch for violations as well, Raub said.

Christopher Anderson

London

As the European Communities (EC) member states debate plans for greater economic and political unity, the European Commission has launched a bold plan to wrest control over the EC's research programme from the Council of Ministers — the periodic meetings of member states' research ministers.

The move would decrease the influence that individual member states have on the precise direction of EC research. This alarms many of the member states' permanent representatives in Brussels, who are embroiled in a dispute with the Commission's controversial vice-president in charge of research, Filippo Pandolfi, over the EC's future support for basic science.

The Commission, the civil service of the EC, is responsible for drawing up the EC's Framework research programmes, in consultation with the member states. But the go-ahead for individual component programmes within each Framework currently depends on the research ministers giving their approval.

The Commission's bid for increased power is put forward as a proposal to the Intergovernmental Conference that is considering plans for greater European unity. The Commission argues that each Framework should be approved in a single step, rather than individual component programmes being subject to scrutiny by the Council of Ministers after the overall budget is agreed. Under the plan, the Commission would then be able to implement the individual programme itself. This would give the Commission a large degree of control over a huge sum of money: a budget of 5,700 million ECU (about £4,000 million) was set last year for the EC's third Framework programme.

The proposal was considered by a meeting of member states' permanent representatives in Brussels late last month. Most are said to have been unenthusiastic. But the Commission has one, superficially appealing, argument to support its plan to streamline the decision-making process: the third Framework programme has become hopelessly bogged down in discussions between the Commission, the European Parliament and the member states. After almost a year, no individual programmes have yet received the research ministers' final approval.

But Pandolfi's critics say that the Commission is to blame for the delay. The problem, they argue, stems from Pandolfi's decision to present the outlines of 13 component programmes within the third Framework at one go. Previously, the European Parliament, the member states' research ministers and their permanent representatives in Brussels were able to consider programmes in smaller, more manageable batches.

To compound the administrative night-

mare, the Commission was last year forced to rewrite the 13 programme outlines, because they contained too little detail. Even now, the "human capital and mobility" programme, which will take over EC support for basic scientific research, is the subject of a bitter row between Pandolfi and the member states' representatives. Pandolfi wishes to spend most of the 518 million ECU (about £350 million) on a large postdoctoral fellowship scheme, which his opponents argue vastly exceeds the demand (see *Nature* 349, 556; 14 February 1991).

In the light of the current chaos, Sir Peter Swinnerton-Dyer, chairman of CODEST, the committee that oversees EC support for basic research, says he cannot see the member states' research ministers agreeing to the Commission's proposal. The European Parliament is also likely to fight the plan. The parliament can influence EC research programmes through its decisions on the EC's annual budget, and is wary of any extension of the Commission's role. With this opposition, the Commission's plan is unlikely to be accepted in its present form. Nevertheless, with the EC's constitution effectively 'up for grabs' as the member states debate greater European unity, it is possible that the Commission may win some concessions and consolidate its role in directing EC research.

Peter Aldhous

UK UNIVERSITIES

New CVCP chairman

London

UK university vice-chancellors have chosen David Harrison from the University of Exeter to head the Committee of Vice-Chancellors and Principals (CVCP) from 1 August. Harrison, a chemist, is the second vice-chancellor to be appointed to the post in recent months. The first, Graeme Davies from the University of



Liverpool, withdrew when he accepted the position of chief executive of the Universities Funding Council (UFC) (see *Nature* 349, 443; 7 February 1991). The two men face a difficult task improving strained relations between their respective bodies.

Peter Aldhous

MITI moves to open patents

Tokyo

FOREIGN companies that join research and development projects funded by Japan's Ministry of International Trade and Industry (MITI) should soon be able to hold patent rights arising out of the research. MITI hopes that the move will encourage foreign companies to take part in research and development projects in Japan.

Until now, all patent rights arising from projects funded directly by MITI have been held by the government, and participating

NEDO and MITI projects free of charge or for a "small fee". The powerful ministry has previously spurred Japan's economic success through aggressive national projects, and the patent legislation is seen as the first step in a new MITI policy of promoting international research on new technology among developed nations.

The new "technoglobalism" policy, which MITI announced at a meeting last year of the Organisation for Economic Cooperation and Development, is designed to counter

PATENT RIGHTS FOR COMPANIES PARTICIPATING IN
GOVERNMENT-FUNDED RESEARCH AND DEVELOPMENT PROJECTS

Country	Holder of rights	Licence fee
Japan		
National project	Now: government Revised: 50 per cent or more government; up to 50 per cent for companies	Fee required Free or small fee
NEDO project	Now: 50 per cent or more NEDO; up to 50 per cent for companies Revised: as above	Fee required Free or small fee
US		
US companies	Companies	Free
Non-US companies	Government	Free
UK	Government	Free
France	Companies	Free
Germany	Companies	Free

companies — both foreign and Japanese — have to pay licence fees to use technology arising from their own research. In the case of projects funded by the New Energy and Industrial Technology Development Organization (NEDO), a semi-government organization affiliated to MITI, companies can hold up to 50 per cent of patent rights, but they still have to pay licence fees.

Under a proposed new law, which MITI submitted last week to the cabinet, companies will have access to patents from

a perceived "technonationalism" in the United States, whereby the United States tries to protect its own technological advances from countries such as Japan that are making rapid progress in developing high technology.

The move to introduce the new legislation was in part spurred by a large-scale project launched by MITI through NEDO in 1989 to develop a "super/hypersonic transport system" to follow the Anglo-French Concorde. US and European companies, including

Rolls-Royce of the United Kingdom, United Technologies, Pratt and Whitney, and General Electric of the United States, and SNECMA of France, wanted to join the project, in which MITI will invest ¥28,000 million (\$200 million) over eight years, but they were discouraged by the regulations governing patent rights. Agreement for these companies to join was finally reached a few weeks ago, after months of wrangling over patent rights.

The legislation will bring Japan more closely into line with France and Germany (see table). In the United States, foreign companies cannot hold patent rights arising out of government-funded research, but they can license technology free of charge if they participate in a project. In the United Kingdom, the government retains patent rights for inventions that arise from contract research, but companies can hold patents when the government only provides subsidies.

MITI officials say they are optimistic that the Science and Technology Agency will soon follow suit with more general legislation covering all Japanese national research projects.

David Swinbanks

PREVENTIVE MEDICINE

AIDS control through cultural heritage

New Delhi

ANCIENT scriptures and religious books will be searched for messages that can be used to control the spread of AIDS in a new programme launched by the Indian Council of Medical Research (ICMR). The council says that Western-style educational campaigns, which often rely on television and printed booklets, will have little effect on Indians, whose life style is shaped more by religion than by science.

A. S. Paintal, director general of ICMR, notes that the United States and African countries have had difficulty controlling the spread of AIDS despite printed information and television advertisements. India has an advantage over Western countries in that it has a rich cultural heritage and scores of scriptures that already contain the necessary prescriptions for an AIDS-free society, Paintal says.

This led to the idea for the campaign, whose theme is "Protect yourself from AIDS through cultural heritage". The hope is that AIDS messages based on religion will be more easily accepted by an average Indian than the same messages given as bland scientific facts through health bulletins. Paintal says that Indian scriptures are replete with messages preaching moderation in sex and describing prostitution, homosexuality and adultery as sin. "All we are saying is, make our people aware of these teachings", Paintal says. "Just telling them about safe sex will not help in changing their attitudes." K.S. Jayaraman

INTERNATIONAL COLLABORATION

Euro-laser plans on ice

London

EUROPEAN science administrators have shelved plans to build a laser facility that would compete with the best the United States and Japan have to offer. The proposed ten-year collaboration among Britain, Germany, France, Italy and Spain was aimed at producing a 100-kilojoule ultraviolet laser that would give European physicists a state-of-the-art machine with which to study plasma behaviour and the fusion of hydrogen isotopes (see *Nature* 346, 503; 1990). But Britain, France and Germany have now decided that the project is too expensive.

The British contribution is among the projects axed to ease the financial difficulties at the Science and Engineering Research Council (SERC, see *Nature* 349, 551; 14 February 1991), and the high costs of reunification are currently restricting

German spending on international science projects. The original proposal would have cost up to £25 million a year and would have required first building two smaller machines to test rival laser technologies before constructing the final 100-kJ facility.

Michael Key, from SERC's Rutherford Appleton Laboratory and one of the authors of the European laser proposal, says the project may be revived when the economic situation across Europe improves. Key takes heart from the example of the European Synchrotron Radiation Facility, now being built in Grenoble, which was approved only after several years of discussion. In the short term, Key says the priority now is to upgrade existing lasers around Europe, and to encourage their shared use, perhaps through funding from the European Communities.

Peter Aldhous

Predicting earthquake risks

San Francisco

FOR the first time, scientists are planning to assemble a comprehensive model of the earthquake hazards for a single earthquake-prone region. Focusing on Southern California and integrating 30 years' worth of data from disparate sources, researchers at the newly established Southern California Earthquake Center (SCEC) hope to improve earthquake forecasting and help to predict the strength of ground motions in specific locations in the Los Angeles area.

In particular, this 'master model' should help emergency planners and urban developers, as well as Earth scientists, to attain a better understanding of exactly what earthquake hazards exist in which areas of Southern California.

Although scientists at scores of institutions study California's frequent tremors, no one has ever attempted comprehensive seismic predictions or maps of hazard zones for the region, says SCEC's executive director, Thomas Henyey. But with \$5 million from the National Science Foundation (NSF) and the United States Geological Survey (USGS) in its first-year budget, and an expected 5–11 years of support as an NSF Science and Technology Center, the centre plans to change that. Located at the University of Southern California, it will draw on scientists from seven universities and the USGS.

The model will undoubtedly require much new research, but adequate data already exist to develop rudimentary hazard analysis maps, Henyey says. And with refinement, the maps may serve very practical purposes.

INTERNATIONAL COLLABORATION

Reforms sought from UK Treasury

London

THE Treasury should protect the UK research councils when the cost of participating in international research programmes escalates owing to fluctuating currency exchange rates, according to the House of Lords Science and Technology (S&T) Committee.

When an international collaboration is based overseas, the British contribution must be paid in the host country's currency. The S&T committee's report on international collaboration in science, published yesterday (13 March), says that the Treasury, rather than the research councils, should pay for cost increases caused by the pound's value dropping by more than two-and-a-half per cent against a pre-determined value. The Lords describe the current situation, in which the research councils must play the international currency markets in an attempt to minimize the impact of exchange rate fluctuation, as "grotesque".

If accepted by the government, the propo-

"Soon, we hope you'll be able to come to us as a developer and ask what is the probability of a quake on a particular property and how strong the ground motion will be", he says. "Our goal is to give you our best description of the quake hazard that will be location-specific."

The database will also be used by the National Earthquake Prediction Evaluation Council to improve both short-term and long-term earthquake predictions, says California state geologist James Davis, a member of the council. In addition, SCEC will probably collaborate with the National Center for Earthquake Engineering Research at the State University of New York in Buffalo, another NSF-funded centre.

Henyey expects the centre to take at least two years to integrate existing earthquake-related data into a useful form. The research areas that will be important to the centre's master model include the history of fault movements, earthquake statistics, patterns of strain buildup and crust deformation, physical properties of earthquake sources and descriptions of surface shaking in response to seismic waves. Once the database is complete, it will be used to help guide the centre's future research directions, but SCEC is already reviewing the nearly 40 research proposals it has received from interested scientists, and it expects to make funding decisions on them in the next few weeks.

The SCEC may have a Northern California counterpart soon, Davis says. Researchers in the San Francisco Bay area are considering development of a similar organization for their region, which, like Southern

sal would ease a long-standing concern of the Science and Engineering Research Council (SERC), which has almost one-third of its annual budget tied up in international projects. The committee notes that SERC was forced to bear the cost of a one-third real-terms rise in the UK subscription to CERN, the European particle physics laboratory in Geneva, between 1985 and 1988.

The S&T committee also attacks the Treasury's practice of reducing the domestic British spending on science to compensate for money that UK scientists win from the European Communities' research programmes, joining a chorus of criticism from other all-party parliamentary committees (see *Nature* 346, 305; 1990 & 349, 183; 1991). This 'attribution' means that, often, only one-third of Britain's spending on EC research represents a real addition to the domestic spending on science. The S&T committee's report notes that neither France nor Germany operates a similar system.

Peter Aldhous

California, is populous, prone to earthquakes and well studied by Earth scientists.

SCEC is set up as a 'center without walls' — an organization of people in many institutions rather than in a single location. In addition to USC and the USGS, the centre's participants include the California Institute of Technology, the University of California at the campuses of Los Angeles, San Diego, Santa Barbara and Santa Cruz, and Columbia University's Lamont-Doherty Geological Observatory in New York.

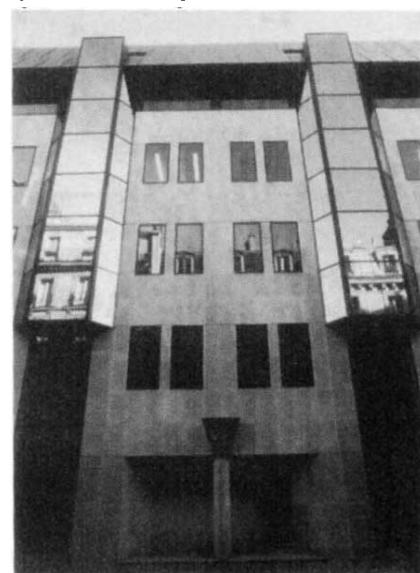
Elizabeth Schaefer

PASTEUR INSTITUTE

New building defies the metro

Paris

THE French Prime Minister, Michel Rocard, last week inaugurated a new retrovirus building at the Pasteur Institute in Paris. Costing FF100 million, the 4,685-square-metre building was paid for entirely by funds raised by the institute since 1987.



Laboratory on springs.

There are ten laboratories with space for 120 researchers and the same number of technicians and administrators, mostly working on AIDS.

The building incorporates some unusual features. The whole structure, which weighs 12,800 tonnes, is supported on 246 massive springs whose function is to absorb vibrations from a Paris metro line that runs just a few metres from the foundations. And to make sure viral material does not escape into the atmosphere, each of the four maximum-security laboratories has total air filtering, special extractors and autoclaves, and is kept below ordinary atmospheric pressure.

Since 1984, spending on AIDS research at the Pasteur has increased by a factor of six and accounts for about one-fifth of the FF600-million annual budget in 1990.

Peter Coles

Complexity still running

SIR — Your correspondents (*Nature* 348, 280; 1990) trying to tackle my question about runaway complexity (*Nature* 347, 704; 1990) have missed the point, which concerns the inverse relationship between morphological complexity and population size. Invoking the eukaryote/prokaryote distinction is not 'the simple answer' as offered by Philip Wood, because the origin of eukaryotes from prokaryotes is itself only part of the more general problem. I accept, of course, the claim of great biochemical versatility, adaptability and general success of today's prokaryotes, made on their behalf by an unusually articulate variant of *E. coli* (*Nature* 349, 97; 1991).

One might loosely define more complex organisms as having more morphological features whose presence is ascribed to the sometime action of directional selection. But more complex organisms generally have smaller population sizes and tend to have longer generation times, so *en route* to increasing complexity there should be fewer and fewer opportunities to build up any new complex feature step-by-step in a sustained series of selection events. Most scenarios to generate the phenomenal complexity of even the simplest metazoans seem to require vast numbers of successive selection events acting on particular traits; that is, episodes of persistent directional selection for one feature after another. The awesome morphological complexity of organisms such as vertebrates that have far fewer individuals on which selection can act therefore remains somewhat puzzling (for me at least), despite the geological timescales available and notwithstanding the insights provided by John Bonner in his book *The Evolution of Complexity* and Richard Dawkins' demonstration of the power of selection in *The Blind Watchmaker*. Could there therefore be a tendency, as yet neither fully recognized nor understood, for more complex organisms to throw up ever wider ranges of viable variation, including occasional novel but well-integrated features that give every impression of sustained selection? If so, any such tendency for 'runaway complexity' may have had a relatively free rein of expression during the metazoan radiations in the late Precambrian and early Cambrian.

In general, we know much about frequency-changing processes in evolution but still very little about how new traits and forms arise, as recently admitted by Endler and McLellan in a comprehensive review (*A. Rev. Ecol. Syst.* 19, 395–421; 1988).

PETER R. SHELDON

Department of Geology,
University of Wales,
PO Box 914,
Cardiff CF1 3YE, UK

Origins of AIDS

SIR — Why do Western analysts need always to invoke exotic sexual practices by Africans in order to explain transmission of HIV-related viruses from monkeys to man when far likelier routes for transmission are readily apparent? In an otherwise excellent letter on the origin and spread of AIDS, A. Karpas (*Nature* 348 578; 1990) cites a paper by F. Noireau (*Lancet* i, 1498; 1987) suggesting that elaborate sexual rituals involving human genitalia and monkey blood are a likely mode of trans-species transmission. Yet in the same letter, Karpas discusses monkey-virus transmission to (Western) laboratory workers. Suggestions of deviant sexual behaviour are no longer evoked — simple monkey bites are now considered sufficient to explain transmission. Man has been hunting, killing and eating monkey in Africa for millennia. This has undoubtedly led to countless bites, cuts and abrasions that have provided ideal opportunities for blood-borne inoculation. Scientific writers should leave descriptions of the more exotic sexual rituals to either anthropologists or the glossier pages of other types of publication, rather than *Nature*.

JOHNJOE MCFADDEN

Department of Microbiology,
University of Surrey,
Guildford, Surrey GU2 5XH, UK

Genome project

SIR — Biologists should be disquieted by the special pleading of Walter Gilbert for the genome project (*Nature* 349, 99; 1991). The assumptions he makes are odd. For instance, he seems to imply that, once we have identified the genes in an organism, we have obtained virtually all the basic information required to explain biological phenomena. The limiting factor is understanding not only the gene structure but also the gene expression, which varies in a most complex way. Monod made great advances towards the understanding of gene expression but only to one case, namely expression of the β -galactosidase gene.

But consider the problem when one becomes seriously ill with, say, a virus infection. Although the genes are the same, a cascade or even catastrophic series of changes occurs in the body functions, or gene expressions, of which man is essentially ignorant. The 'new paradigm' should be addressed not merely to gene structure but to elucidation of the integrated functions of genes, enzymes, membranes and metabolites in the whole organism. Only such an extended paradigm can lead to explanations of the biological phenomena that Gilbert lists in his preamble. Incidentally, formulating a hypothesis and testing it experimentally is not due to start only when we have identified the genes. A review of the literature will show that biologists have always effectively used that principle.

The use of 'kits' and computers on the grand scale for the human genome project is an admirable ambition. But let us recognize its limitations and ensure that it does not distort the rest of bioscientific research. The attack on the gene expression problem is much more challenging and difficult. The resort to the kit and the computer routines seems to be an example of that well-known human failing — taking the line of least resistance.

S. JOHN PIRT

(Emeritus professor,
King's College London)

50 Chartfield Avenue,
London SW15 6HG, UK

Grants as status

SIR — Ray Iles (*Nature* 349, 188; 1991) is lucky to be able to court the biotechnology companies. Obvious sources of funding in anaesthesia are the pharmaceutical companies — except that their money does not count in the scheme of 'brownie points' erected by those who judge our worth, to whom the acquisition of grant money seems the most important criterion. However, in these days of shrinking resources one would imagine that doing research for the least cost would be a laudable aim; yet one scores more highly by securing grant money, even if the project can be done without its need. It is an intellectually bankrupt notion to suppose that research that is more expensive is necessarily better. I pity the theoreticians; or perhaps they secure grants for supplies of expensive fountain pens and pure vellum on which to write.

NEVILLE W. GOODMAN

University Department of Anaesthesia,
Southmead Hospital,
Bristol, BS10 5NB, UK

Ordered layers

SIR — The intriguing piece by Daedalus on ambitious chemists and micromagnetics (*Nature* 348, 588; 1990) came in the middle of a related discussion here. I have supposed that, if the vapour phase deposition of organic compounds (such as porphyrins, phthalocyanines and other aromatics) which are able to sustain a sizeable induced ring current were to be carried out in a high magnetic field, then the induced magnetic field would influence the mode of deposition. With care it might be possible to get nicely ordered layers in this way; and clearly the experiment could be done with the external field at various orientations to the direction of gaseous molecular flow.

Has this sort of experiment been done? If not, would anyone be prepared to do it? If not in the United Kingdom, then maybe in Japan? I would be happy to supply suitable porphyrins and phthalocyanines.

R. BONNETT

Department of Chemistry,
Queen Mary and Westfield College,
Mile End Road, London E1 4NS, UK

Classical liquid in a box

A simulation of a fluid of interacting particles in a box much narrower in one dimension than the other two reveals a curious bimodal density distribution along the short dimension — and provides a telling example of Poisson statistics.

THAT the perennial question of whether fashions for problems of a particular kind precede or follow the arrival of means for their solution will never be settled is nicely illustrated by the recent interest in smallish clusters of atoms or molecules. The recognition of small clusters in atomic or molecular beams almost coincided with the point at which the techniques of molecular physics could be turned in that direction with some chance of success.

That, no doubt, is one reason why the problems of molecular systems whose dimensions are essentially atomic are now all the rage. But that does not diminish their high interest, which is that of throwing light on the transition between the properties of a collection of isolated molecules and a bulk aggregate thereof. But now there is a model calculation of a system of this kind, essentially a bulk liquid consisting of classical atoms, in which only one of the three physical dimensions is of an atomic scale. Its outcome is full of surprise.

The problem, definable theoretically if unattainable in the machine-shop, is that of a liquid enclosed between infinite parallel plates which are separated from each other by only three atomic diameters. It is the problem of a fluid in a flattened capillary tube. There are obvious connections with that of a thin layer of fluid adsorbed on the surface of a solid as a two-dimensional gas (but the constraints are greater), as well as with the behaviour of a fluid in a porous structure (but "micropore" would be a better word).

J. G. Powles, A. Dornford-Smith and W. A. B. Evans from the University of Kent at Canterbury (England), who have a reputation in the solution of static versions of problems such as this, now say that their interest is in the dynamics of atomic motion in such a system (*Phys. Rev. Lett.* **66**, 1177–1180; 4 March 1991).

Describing the problem is easier than solving it. Pairs of atoms are supposed to interact by means of a Lennard–Jones potential — the standard combination of an attractive force varying with the sixth power of the inverse separation and a repulsive force varying similarly with the twelfth power (which dominates at short distances). The 'walls', simulated as an external potential applicable to all atoms between them, repel all atoms according to the ninth power of their distance, thus preventing what might be called condensation on the surface.

The bare bones of static equilibrium are

easily guessed at. If the fluid were a perfect gas, so that the mutual interaction of pairs of molecules vanishes, but if the repulsion of the walls persists, there will be a smooth distribution of molecules within the slab of space between the walls, with maximum density midway between them. Increasing temperature (or giving the atoms kinetic energy) makes the distribution flatter. But what happens if the mutual interaction is switched on? Then there is a relative concentration of molecules in two layers on either side of the midpoint (which is only, in this model, 1.5 atomic diameters from each wall).

Handwaving will get one almost there. Atoms are repelled both by the walls and by those of their own number lying closer than just over one atomic diameter. (The energy minimum of the Lennard–Jones potential is at 1.12 times the number taken to be the diameter.) So a density distribution in which atoms were crowded near the midpoint would be unstable relative to a bimodal distribution, but less so at increased temperature.

In practice, these properties are arrived at by numerical calculation. Powles, Dornford-Smith and Evans put 1,020 atoms in a simulated box and follow what happens to them. As it happens, the two long dimensions of the box turn out to be just over 23 atomic diameters — far from infinite, but probably sufficient for the purposes. It turns out that there is little of interest to say about the motion of atoms parallel with the confining walls; even if one of them should embark from a place where the density is a maximum, only rarely will it thread its way through the relatively dense gas at that exact distance from one wall. So classical diffusion is a sufficient picture of what happens.

The fun comes in the direction perpendicular to the wall. What emerges from individual trajectories is that atoms spend the vast majority of the time at or near one or other of the two density maxima between the walls, transferring from one extremum to the other at high speed. At any time, only about ten per cent of the atoms will be in intermediate positions; in other words, the transition time is about a tenth of the average time spent at one of the extrema.

So much has become apparent largely because the numerical calculations consist of iterations of the state of the system at intervals of time that are extremely short compared with the physical characteristics of the system. The basic time-step used, for example, is equivalent to one twentieth of the

velocity auto-correlation time, which is an inverse measure of the time over which velocities persist unchanged. A coarser time-step might well have obscured this fine detail.

In itself, this says little for the behaviour of real fluids tightly constrained by external boundaries except that they are unlikely to behave as fluids in the directions in which they are constrained, whatever their properties in perpendicular directions. On the face of things, the transverse behaviour of the fluid involving the hopping of atoms from one favoured position to another may be a sign that the fluid behaves as a liquid in that direction and as a gas in the two other directions.

It would be interesting to see what would happen to this model (as to other models) if the basic parameters were different, and in particular if they were such as to make condensation probable and recognizable (which is not that simple in such a narrow slab of fluid). Whether it would be possible to extend the calculations to atoms with directional preferences, as might be expected of water molecules, is more problematical. Already, the authors note that each time-step of their calculation takes 1.8 seconds of time on an array of 17 parallel-wired transputers — the British-made microprocessors that come nearest to being computers on a single chip. That seems to imply a few days for each simulation.

But Powles, Dornford-Smith and Evans seem most of all intrigued by having stumbled on a neat way of testing whether experimental data fit Poisson statistics. Modishly, they refer to an article dating from 1910 in which Rutherford, Geiger and Bateman regret the difficulty of demonstrating the Poisson nature of radioactive decay by accumulating data about the intervals of time between successive decay-events (*Phil. Mag.* **20**, 698; 1910), chiefly for lack of data-handling techniques.

The case of the constrained fluid is more easily handled. If an event is the crossing of an atom from one extremum to the other, the intervals, t , between events are readily calculated and recorded, from which it is possible to work out all possible quantities such as $\langle t^n \rangle$, where n is an integer and the angle-brackets mean averages. But the outstanding characteristic of a Poisson distribution is that $\langle t^n \rangle / \langle t \rangle^n = n!$, whatever the particular parameters defining the distribution. So the constrained fluid is christened a "liquid Poisson oscillator". The name is inconsequential, but the model is not.

John Maddox

Seeds of a universal tree

David Penny and Charles J. O'Kelly

ACCORDING to the serial endosymbiosis theory, mitochondria and chloroplasts arose by intracellular symbioses between ancestral eukaryotes and free-living unicellular organisms. For these organelles the principle is now well established. But are they monophyletic or polyphyletic, and if polyphyletic,

elles within this membrane system could be the remains of an endosymbiotic eukaryote, with the nucleomorph being its vestigial nucleus. If this were true, then *Cryptomonas* would be expected to have two sets of eukaryotic ribosomal sequences.

From total *Cryptomonas* DNA Douglas *et al.*¹ were indeed able to isolate and sequence two functional 18S RNA molecules, whereas nuclear DNA yielded only one of them. The inference that the other is from the nucleomorph is reasonable. The sequences were analysed by distance matrix and maximum parsimony methods and added to the now familiar, ever-growing universal evolutionary tree.

It is always important to estimate the reliability of an evolutionary tree. To do this, Douglas *et al.* used bootstrap resampling methods to generate new sets of data and inferred evolutionary trees from each set. The resulting trees are compared to find the most stable branches on the original tree. Bootstrapping is now a minimal requirement for an evolutionary study. In the present case, assignments within some major eukaryote clades seem to be robust although relationships between major clades are not. But importantly, the cryptomonad nucleus and nucleomorph appear on different places on the tree and their relative separation seems to be robust. The nuclear rRNA is most closely related to that of fungi, *Acanthamoeba* and green plants, whereas the nucleomorph rRNA segregates with nuclear rRNAs from red algae.

The phylogenetic position of the nucleomorph sequence is supported by other molecular data. For example, both large and small subunits of the photosynthetic enzyme ribulose biphosphate carboxylase are encoded on cryptomonad chloroplast DNA³; only the large subunit is encoded on the chloroplast DNA of plastids containing chlorophyll *b*. In addition, sequences of both subunits of this enzyme in *Cryptomonas* Φ plastids are most closely related to those of red algae. Indeed, the cryptomonad and red algal sequences form a clade separate from the one that includes cyanobacteria, cyanelles and green plants^{3,4}. Both the rRNA and ribulose biphosphate carboxylase gene sequences indicate separate origins at least for 'green' and 'red' plastids, and that red plastids have been secondarily acquired by cryptomonads.

Another important conclusion comes

from comparing *Cryptomonas* nuclear sequences with those of other eukaryotes. Cavalier-Smith⁵ postulated that cryptomonads belong to a monophyletic kingdom Chromista, the plastid-bearing members of which have a chloroplast endoplasmic reticulum membrane system. On this model the chloroplasts of cryptomonads arose from a single endosymbiotic event in the ancestral 'eochromist'. Accordingly, nuclear rRNA sequences from cryptomonads and chromists should occur together on the tree. But Douglas *et al.*¹ find that the *Cryptomonas* nuclear sequence is separated by a considerable phylogenetic distance from chromist sequences. Thus the rRNA sequences indicate that chloroplasts in these two groups have arisen from separate endosymbiotic events.

These bits of molecular evidence support predictions, based initially on structural and biochemical studies, that chloroplasts have arisen many times. How do we explain this result? Proponents of chloroplast polyphyly have drawn most of their evidence from observations on plastid diversity, and their ideas are, in essence, *ad hoc* explanations. On the other hand, proponents of monophyly have argued on the basis of the mechanisms involved in the transformation of an endosymbiont into an organelle. This transformation requires many steps, each apparently improbable. For example, it is improbable that when a chloroplast gene is transferred to the nucleus, its protein product will have the correct signals to be transported back into the chloroplast. Cavalier-Smith has therefore argued, logically and consistently, for phylogenies that minimize the number of endosymbiotic events.

The past should not be 'explained' in terms of highly improbable events. But as more data became available the predictions of monophyly, starting from no endosymbiotic events, have had to be modified. How has the apparently impeccable reasoning of the proponents of monophyly misled us?

Part of the answer may lie in mechanisms of transformation still to be discovered. For example, transport of phycoerythrin across thylakoid membranes in *Cryptomonas* Φ does not seem to involve a normal signal peptide⁶. But the answer must also partly be in the other side of the equation — the many opportunities for organelles to form during

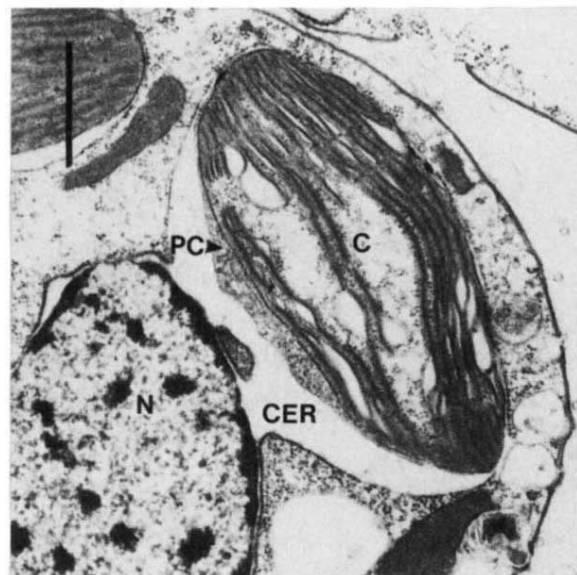


FIG. 1 Cross-section of *Cryptomonas* photosynthetic compartment showing the chloroplast endoplasmic reticulum (CER). The CER contains the nucleomorph (NM), chloroplast (C) and periplastidial cytoplasm (PC). The outer membrane of the CER is continuous with the nuclear envelope and is studded with ribosomes. N, nucleus. Chromophytes have a similar CER but lack the nucleomorph and periplastidial cytoplasm. Chlorarachniophytes possess the nucleomorph and periplastidial cytoplasm, but the outer 'CER' membrane lacks ribosomes and continuity with the nuclear envelope. Scale bar, 1.0 μ m.

from how many independent ancestral lines do they derive? On page 148 of this issue¹, Douglas *et al.* now report on an analysis of small-subunit (18S) ribosomal RNA sequences from a cryptomonad flagellate (*Cryptomonas* Φ) which shows that chloroplasts have polyphyletic origins, and that some eukaryotes obtained their chloroplasts from other eukaryotes.

Why study cryptomonads? This aptly named group of phytoflagellates has a perplexing assortment of structural and biochemical features², rather as if they had browsed among the other groups of protists and picked out the features that they liked. Amid the confusion, one structure is particularly notable: the nucleomorph. This organelle, bound by a double membrane, looks like a nucleus on a diet and contains both DNA and RNA. What is more, together with the chloroplast(s) and associated cytoplasm, the nucleomorph is enclosed within the 'chloroplast endoplasmic reticulum' membrane system (Figs 1 and 2). Even when the nucleomorph was first described some 25 years ago, it was recognized that the organ-



FIG. 2 The *Cryptomonas* nucleomorph (NM) with double membrane and nucleolus-like body (asterisk) lies within the CER next to the chloroplast (C). PC, periplastidial cytoplasm. The nucleomorph is about 1 μ m in diameter at its widest point and is distinctly flattened; the nucleus is 3–5 μ m in diameter at its widest point and is more nearly spherical. Scale bar, 0.5 μ m. Micrographs by C. J. O. and DSIR Electron Microscopy Unit (Palmerston North).

the lengthy course of eukaryote evolutionary history. An analogy is the recent observation⁷ that although only one random RNA molecule in 10^{10} may bind a specific dye, if a mole of random RNA molecules (roughly 6×10^{23} molecules) is made, then this vanishingly rare event becomes a certainty.

Let us assume that the same ecological principles applied to the early eukaryotes as apply today. If so, most early photosynthetic organisms would have been protected by a multitude of endo- and exo-symbiotic arrangements⁸. Assume on average that there was the equivalent of a single layer of photosynthetic symbionts (a 'symbiont area index' of one) over the surface of the earth ($510 \times 10^3 \text{ km}^2$), each with a surface area of $10 \mu\text{m}^2$. This gives the approximate number of photosynthetic genomes as about 10^{20} . If each divided ten times a year then after 100 million years 10^{29} divisions would have occurred, each offering an opportunity for gene transfer between symbiont and host. If anything, this crude calculation is an underestimate of the number of genomes available for transformation into plastids. Success may be rare but opportunity is great.

It would be a mistake, though, to believe that we have done more than begin a proper inquiry into organelle polyphyly and the origin and early diversification of eukaryotes. In much of the present work, too much reliance is placed on results from a single gene, and too much other information is ignored. Study of groups, such as mammals, that diverged much more recently than protists indicates that it is necessary to combine the analyses of several sequences to get results with any sort of reliability⁹. It would be surprising to find that a single gene, even a long one, were sufficient to infer robust trees for divergences that occurred a billion or more years ago. Even in single-gene sequence analyses, additional data are often needed before the trees can be considered to be reliable, as the bootstrapping results by Douglas *et al.*¹ indicate for 18S rRNA trees.

As a result, several fundamental questions remain. At least some chromist chloroplasts, according to sequence analysis of ribulose biphosphate carboxylase genes, belong in the same clade as cryptophyte and red algal chloroplasts³. The resulting dilemma seems to be best resolved by proposing that the endosymbionts separately obtained by cryptophytes and chromists were closely related red algae. But will this explanation survive more rigorous analysis? Differing G+C nucleotide

contents seem to confound the relationship between prochlorophytes and green plant plastids⁸ and analyses will have to adjust for this. More rigorous analysis would also help confirm that green plant plastids are related to glaucocystophyte cyanelles but originated separately³.

The most immediate progress in this research area will probably come from studies designed around a specific evolutionary question, often suggested from the literature of protist structure and biochemistry, and

MOLECULAR EVOLUTION

The resistance movement

J. F. Y. Brookfield

ATTEMPTS to reduce the incidence in developing countries of diseases transmitted by insect vectors have frequently been hampered by the evolution of insecticide resistance in the vector species. One of the best examples of this problem comes from the efforts to quell the spread of disease by mosquitoes. Mosquitoes of the *Culex pipiens* complex often become resistant to organophosphorus insecticides. Resistance arises from the overproduction of nonspecific esterase enzymes which is usually mediated by the amplification of the corresponding structural genes¹.

Raymond *et al.* now report on page 151 of this issue² that the restriction maps of the amplified B2 alleles of the esterase gene in mosquitoes from around the world are indistinguishable. The authors conclude that the amplified B2 genes probably all originated from an initial amplification event that spread rapidly throughout the world by migration. As well as having implications for those interested in designing insecticides against mosquitoes, the new finding shows just how informative DNA sequences are in reconstructing the evolutionary history of alleles within populations.

The allelic esterase alleles B1 and B2 are often amplified as many as 250 and 60 times respectively in mosquito strains which overproduce esterase. Raymond *et al.*² analysed the restriction maps of amplified B2 esterase genes in strains from Texas, California, Pakistan, the Ivory Coast, the Congo and Egypt, and found that the maps were the same for the 24 restriction sites at this locus. An experiment of this kind not only looks directly at the 144 bases found at the restriction sites, but also looks indirectly at other six-base sequences that differ by a single base substitution from a restriction site. Furthermore, the authors found that the sequences do not differ by any insertions or deletions of DNA which are frequently found among flanking sequences of *Drosophila* alleles³. Although it is impossible to calculate precisely the statistical strength of the inference, the results imply that all these amplified alleles had a recent common ancestor. The most likely explanation is that a single B2 allele was amplified and that the resulting amplified array

will involve thorough analysis of not one but several molecular sequences and checking of the results against theories generated by the non-molecular disciplines. The body of work that is accumulating on *Cryptomonas* Φ has many of the required attributes, and suggests that the simplistic first stage of the search for the single universal evolutionary tree is coming to an end. □

David Penny and Charles J. O'Kelly are in the Department of Botany and Zoology, Massey University, Palmerston North, New Zealand.

spread throughout the world owing to the selective advantage that it confers under insecticide treatment.

The result is of great significance to those interested in designing schemes for maximizing the effectiveness of insecticides used against mosquito populations. If insecticide-resistance alleles can become ubiquitous throughout the world by migration from where they initially arise, rather than through their independent occurrence in different places, then other resistance alleles found worldwide could result from rare events in the insect genome. Small-scale trials of new insecticides over limited areas may thus compromise their subsequent large-scale use, because new resistance alleles selected in small areas may quickly spread throughout the range of the species.

The work also has a more general importance for those interested in adaptive changes in wild populations. First, it shows that there can be a remarkably rapid spread of a new allele across continents in mosquitoes, probably by the unintentional help of humans⁴. Second, it illustrates another way in which the information content of DNA can be used to reconstruct the history of changes in allelic frequencies in wild populations. There is a long history of study of genetic polymorphisms defined by charge differences in soluble enzymes⁵. Attempts to understand the selective forces, if any, acting on such polymorphisms have often been confounded by the possibility of heterogeneity existing among alleles belonging to the same electrophoretic class. Two alleles indistinguishable by the electrophoresis of their products may not be identical at the DNA level and thus may not necessarily be subject to the same selection⁶. Nor would their apparent identity necessarily result from shared descent.

The analysis of DNA sequences, however, has held out much hope for those wishing to discover the evolutionary history of alleles, partly because of the high information content of DNA sequences. In *Drosophila*, where many alleles have been surveyed, the heterozygosity at the base-pair level is, as would be expected, variable between loci; and estimates based on restriction-map vari-

1. Douglas, S.E., Murphy, C.A., Spencer, D.F. & Gray, M.W. *Nature* **350**, 148–151 (1991).
2. Gillott, M. in *Handbook of Protozoa* (ed. Margulis, L., Corliss, J.O., Melkonian, M. & Chapman, D.J.) 139–151 (Jones & Bartlett, Boston, 1990).
3. Valentin, K. & Zetsche, K. *Pl. molec. biol.* **15**, 575–584 (1990).
4. Douglas, S.E., Durnford, D.G. & Morden, C.W. *J. Phycol.* **26**, 500–508 (1990).
5. Cavalier-Smith, T. *Prog. phycol. Res.* **4**, 309–347 (1986).
6. Reith, M. & Douglas, S. *Pl. molec. Biol.* **15**, 585–592 (1990).
7. Ellington, A.D. & Szostak, J.W. *Nature* **346**, 818 (1990).
8. Lockhart, P. thesis, Univ. Sydney (1990).
9. Penny, D., Hendy, M.D., Zimmer, E.A. & Hamby, R.K. *Aust. Syst. Bot.* **3**, 21–38 (1990).

ation generally fall in the range from 0.2 to 1.2 per cent⁷⁻⁹. This means that even at the level of restriction mapping, large numbers of distinguishable alleles exist at a typical locus, and with the use of the polymerase chain reaction to amplify and sequence stretches of DNA yet more allelic variants will be detected.

The other advantage that DNA sequences offer in tracing the evolutionary history of alleles is that most base differences between alleles probably have very little effect on the phenotype. Thus most changes in DNA sequence do not result from natural selection, but occur randomly. As a result, there is little chance of such changes being convergent. It is this assumption that allows Raymond *et al.* to conclude from the identity of their restriction maps that the amplified B2 alleles arise from recent shared descent.

Studies of DNA sequences may thus considerably increase the precision with which the evolutionary history of genes within populations can be reconstructed. It has long been known that the DNA sequences of a particular gene from different species can be used to construct an estimate of the phylogeny of the alleles, and thus of the species possessing them. Similarly, the genetic variation between alleles within species is so great that it is possible to study the phylogeny of the alleles within a population. This phylogeny is, itself, the consequence of the selection and genetic drift that have occurred at the locus. And in turn, comparison of the allelic phylogeny with expectations based on the neutral theory can indicate that selection has taken place.

The phylogeny of a collection of alleles from a population is not a well-defined concept, of course, because recombination causes such phylogenies to vary between different parts of a gene. This, indeed, forms the basis of the power of the technique, as it is the differences between the phylogenies connecting the alleles of different parts of a gene that give clues as to the nature of the selection operating. Thus because vertebrate mitochondrial DNAs do not recombine, the whole molecule must share the same evolutionary history. Consequently, if selection has played a part in the determination of the phylogeny of mitochondrial DNA, there is no way of knowing which of its genes have been of particular selective importance. □

J. F. Y. Brookfield is in the Department of Genetics, School of Biological Sciences, University of Nottingham, Nottingham NG7 2UH, UK.

1. Mouches, C. *et al.* *Science* **233**, 778–780 (1986).
2. Raymond, M., Callaghan, A., Fort, P. & Pasteur, N. *Nature* **350**, 151–153 (1991).
3. Aquadro, C. F. *et al.* *Genetics* **114**, 1165–1190 (1986).
4. Andreadis, T. G. *Am. Mosq. Control Ass.* **4**, 256 (1988).
5. Selander, in *Molecular Evolution* (ed. F. J. Ayala) 21–45 (Sinauer, Sunderland, 1976).
6. Coyne, J. A., Felton, A. A. & Lewontin, R. C. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5090–5093 (1978).
7. Eanes, W. F., Labate, J. & Ajioke, J. W. *Molec. Biol. Evol.* **6**, 492–502 (1989).
8. Langley, C. H. & Aquadro, C. F. *Molec. Biol. Evol.* **4**, 651–663 (1987).
9. Leigh Brown, A. J. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5350–5354 (1983).

Great balls of fire

Stanley Singer

KAPITZA proposed in a classic paper that ball lightning is formed by short radiowaves generated in natural atmospheric processes¹. Now, extending experimental studies by previous investigators of radiofrequency discharges, Ohtsuki and Ofuruton, who report on page 139 of this issue², have produced long-lived fireballs in air. The motion of the discharges against the flow of gas and their apparent penetration through a dielectric surface are of particular interest, having been among the most difficult to explain of the properties observed in natural ball lightning.

Ball lightning has been the subject of investigation by distinguished figures in science since the early nineteenth century. Although its appearance is very haphazard and far less frequent than ordinary linear lightning, it is a source of astonishment when the ball of fire is observed in the lower atmosphere, often entering dwellings while floating at a leisurely pace. By its random appearances, it has eluded measurement with the scientific instruments which have provided so much information about ordinary lightning. Even photographs are rare. However, from the qualitative reports of eye-witnesses (now numbering in the thousands, particularly from collections from the Soviet Union) the general properties can be deduced despite the wide variability found in the reports.

Scepticism as to the actual existence of ball lightning has occasionally been based on the assertion that no qualified scientist has ever observed the fireball. The assertion has been incorrect for almost a century. More recently in addition, R. C. Jennison, the noted radioastronomer, and, indeed, Ohtsuki himself have been inadvertent observers of the mysterious fireball; and B. Pippard reported its appearance at the Cavendish Laboratory³.

Ball lightning makes its appearance as a flaming globe, typically orange, red–orange or intense white, less often blue, green or yellow. The ball is usually 25–30 cm in diameter, with much smaller and much larger dimensions noted in individual cases (from 1 cm to 10 m and larger). Occasionally, the balls move against the direction of prevailing winds and penetrate window panes without making a hole in the glass; it is these properties that Ohtsuki and Ofuruton now report from their experimental discharges. An absence of warmth has been noted in some incidents in which the ball passed particularly close to an observer; in others the

witness was burned, or combustible materials were ignited. Occasionally the globe disappears with a frightful explosion.

There is an even greater disparity in the reported conditions for the appearance of ball lightning. For a century and a half, the observations in Europe and the New World were almost invariably associated with thunderstorms and often directly with the

Caught unawares — ball lightning at Gorges de Loup.

occurrence of ordinary lightning. In several cases ball lightning has been seen at the tip of a zig-zag lightning flash or alongside the bright channel of the discharge. There were only a few reports of ball lightning appearing in fair weather or in the absence of a storm. A recent collection of hundreds of observations in Japan, however, contains ten times as many cases in fair weather as in rain or storm⁴. These data render theoretical descriptions of generation of the fireballs from lightning channels somewhat less convincing, at least as the only possible mechanism.

The difficult and long-standing problem of ball lightning has attracted few meteorologists or atmospheric scientists. Rather, physicists constitute the larger part of the company studying the atmospheric fireballs. Numerous theories have been put forward, covering chemical, electrical, nuclear and relativistic models. New states of matter have been proposed to explain the unusual properties reported for ball lightning. Often specialists conclude that solutions must be found outside their own field. Thus, Faraday commented that ball lightning cannot be related to lightning or to any form of electrical discharge because the discharges he was familiar with, exist only briefly and travel at

1. Kapitza, P. L. *Dokl. Akad. Nauk SSSR* **101**, 245 (1955).
2. Ohtsuki, Y. H. & Ofuruton, H. *Nature* **350**, 139–141 (1991).
3. Pippard, B. *Nature* **298**, 702 (1982).
4. Ohtsuki, Y. H. *Baryon* **1** (12), 96–97 (1986).
5. Smirnov, B. M. *Phys. Rep.* **188**, 1–78 (1990).
6. Zheng, X.-H. *Phys. Lett. A* **148**, 463–469 (1990).
7. Muldrew, D. B. *Geophys. Res. Lett.* **17**, 2277 (1990).
8. Lushnikov, A. A., Negin, A. E. & Pakhomov, A. V. *Chem. Phys. Lett.* **175**, 138–142 (1990).
9. Cawood, W. & Patterson, H. S. *Nature* **128**, 150 (1931).

In other recent theoretical work, analyses have been made of spherical fireballs containing internal electromagnetic radiation trapped by a surface plasma layer^{6,7}. The trapped radiation provides the continuing energy needed to maintain the sphere for the

The structure and efficiency of the pigment bed supplying the reaction centre must be optimized by specific positioning of pigments within pigment protein complexes of the photosynthetic membrane. In higher plants the main components of the pigment bed are the chlorophyll *a/b* binding (CAB) proteins. From secondary structure analysis CAB polypeptides are predicted to have two to four membrane-spanning helices, although a three-helix model is widely accepted⁵. Kühlbrandt and Wang identify three helices. But this highlights a problem with some of the putative membrane-spanning helices in CAB proteins, namely that it is impossible to model three helices without including charged residues in the membrane (see figure). The CAB helices are also interesting in that two are considerably

109

longer than expected or required simply to cross the hydrophobic part of the membrane, and are longer than those found in the reaction centre or bacteriorhodopsin, which may indicate some anchoring or binding function outside the membrane.

Another question which must be answered for light-harvesting complexes has to do with where the chlorophylls are bound. Crystals of bacterial light-harvesting proteins contain only 18 chlorophylls in an oligomer of relative molecular mass 84,000 (M_r 84K)⁶, whereas the CAB proteins contain 15 in a polypeptide of M_r only 25K. Chlorophyll appears to be associated with specific residues within reaction centres, the 'traditional' residues being histidine, glutamine and asparagine. But there are not enough highly conserved residues among the family of CAB proteins to bind the associated chlorophylls. Unless there are specific residues in each type of CAB, this implies that there is some other form of liganding through water molecules of backbone carbonyls. In the structure determined by Kühlbrandt and Wang, the chlorophylls are arranged on the outside of the protein, which may perhaps facilitate energy transfer to the reaction centre or between CAB molecules. Combination of the results from further refinements of the present structure, with results from X-ray analysis of three-dimensional crystals of the light-harvesting protein from purple photosynthetic bacteria⁶, will allow the development of a full model of light-harvesting structures.

As is only to be expected, the partial resolution of a protein structure poses many more questions than it answers. Inevitably we must ask for even greater resolution and detail, and we need to distinguish the chlorophyll *a* and *b* molecules and define their absorption wavelengths. From the present structure it would seem that there is heterogeneity between the parts of the protein in each leaflet of the lipid bilayer. Does this mean that there is a particular pathway in or out of the monomeric complex to other similar complexes or to the reaction centre? Each plant species contains a number of CAB proteins, and although the structure determined by Kühlbrandt and Wang seems to offer a general model there must be small but important differences between them. The most important question will be how these proteins interact with the reaction centre to provide effective energy transfer. □

M. C. W. Evans and J. H. A. Nugent are in the Department of Biology, University College London, Gower Street, London WC1E 6BT, UK.

Not with a whimper

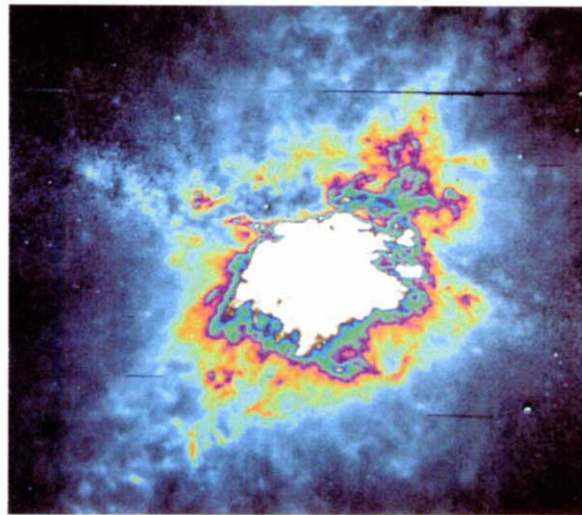
Leo Blitz

THE centre of the Milky Way appears to be primed to explode in a profusion of star formations leading to the expulsion of vast amounts of matter from our Galaxy. This is the conclusion of A. Stark and colleagues, based on an analysis of radioastronomical observations of the central regions of the Milky Way and reported in the *Monthly Notices of the Royal Astronomical Society* (248, 14p – 17p; 1991). This starburst will

tion from the O stars is so intense that it destroys (by dissociation and ionization) the molecular clouds out of which they are born in a sort of cosmic matricide. Finally spent in the relatively short time of 3 million years, an O star suffers a cataclysmic explosion visible as a supernova which leaves a remnant neutron star or black hole.

In our region of the Milky Way, giant molecular clouds typically form fewer than

10 O stars at any one time, and although the phenomena associated with these stars are quite spectacular, their small numbers ensure that their effects are limited to a relatively small volume near the stars themselves. In the centres of some galaxies, however, the star formation is so vigorous that as many as a million O stars are formed almost simultaneously. The collective effects of these stars can produce a brilliant galactic nucleus and give rise to spectacular ejections of matter from stellar winds and supernova explosions. These rare galaxies, known as starburst galaxies, contain unusually large concentrations of molecular gas at their centre. The nearest example of a starburst, shown in the photograph, is the galaxy M82. The bright nucleus is hidden by the intervening dust, but the



The prototypical starburst galaxy, M82. The image shows an extensive system of filaments extending several kiloparsecs above and below the starbursting disk. The kinematics and other physical properties show that gas is flowing out in a large-scale bipolar pattern, probably driven by the starburst. The colours in this false-colour image indicate the intensity of line emission (Balmer- α) from atomic hydrogen. (Courtesy of T. Heckman and L. Armus.)

occur 10–100 million years hence, they suggest, a time very short compared to the age of either the Milky Way or the Solar System.

The Milky Way has been known for some time to harbour massive clouds of gas and dust in its central regions. Similar but smaller clouds occur throughout the disk of the Galaxy, and are the material from which new stars form. Known as giant molecular clouds, these objects are composed primarily of molecular rather than atomic hydrogen, and are typically a hundred to a thousand times denser than the mean for interstellar gas. In our part of the Galaxy, about half way out in the disk, giant molecular clouds typically have enough material to form about 100,000 stars with a mass like that of the Sun. The efficiency is known to be low, however, and only a few per cent of the gas is actually turned into stars.

Occasionally, a few stars form, known as O stars, that are so bright that the radiation from their surfaces accelerates a wind of ions at velocities up to about 1 per cent of the speed of light. The Sun also generates a wind of ions, but at rate as much as 10^{10} times weaker than the O star. The ultraviolet radia-

tion from the O stars is so intense that it destroys (by dissociation and ionization) the molecular clouds out of which they are born in a sort of cosmic matricide. Finally spent in the relatively short time of 3 million years, an O star suffers a cataclysmic explosion visible as a supernova which leaves a remnant neutron star or black hole.

disturbed appearance testifies to the enormously energetic activity going on at the centre. In spite of its spectacular appearance, as starbursts go, M82 is rather wimpy. It has been known since the 1970s that the nucleus of the Milky Way contains a vast reservoir of molecular gas — enough to make more than 10 million O stars if all of the gas could be converted this way. But until recently, there was no reason for supposing that this gas could not maintain itself in stable orbits around the centre, or at worst spiral slowly and harmlessly into the centre owing to the viscosity of the gas. What Stark *et al.* have shown is that the molecular gas that is collected into the largest discrete clouds, which are a substantial fraction of the total, will inevitably spiral into the nucleus in a time no more than about 1 per cent of the age of the Milky Way, and perhaps considerably less.

The reason for this is dynamical friction, a gravitational analogue of what happens if one suddenly shuts off the motor of a boat ploughing through the water. Most of the stars in the nucleus of the Milky Way are part of the central bulge which contains stars ro-

1. Deisenhofer, J., Epp, O., Miki, K., Huber, R. & Michel, H. *Nature* **318**, 618–624 (1985).
2. Feher, G., Allen, J.P., Okamura, M.Y. & Rees, D.C. (1989) *Nature* **339**, 111–116 (1989).
3. Kühlbrandt, W. & Wang, D.N. *Nature* **350**, 130–134 (1991).
4. Henderson, R. *et al.* *J. molec. Biol.* **213**, 899–929 (1990).
5. Karlin-Neumann, G.A., Kohorn, B.D., Thornber, J.P. & Tobin, E.M. *J. molec. appl. Gen.* **3**, 45–61 (1985).
6. Papiz, M.Z. *et al.* *J. molec. Biol.* **209**, 833–835 (1989).

BONE MARROW TRANSPLANTATION

Making just the right match

Peter Parham

tating relatively slowly about the centre. The molecular clouds in the innermost regions of the Galaxy rotate about the centre much more rapidly. Thus, as the molecular clouds move through the more sluggish stellar system, the stars exert a gravitational drag causing the clouds to give up some of their orbital momentum and consequently fall into the nucleus of the Galaxy. The orbital energy of the clouds is dissipated in stellar wakes which increase the energy of the stars, just as the hull of a boat raises a wake on the surface of the water. As the clouds fall into the centre they will collide and coalesce, forming massive aggregations which are ripe for the formation of the vast number of O stars inferred to be present in starburst nuclei.

It is likely that an effect considered by Stark *et al.*, but not explicitly analysed by them, will cause the infall of the molecular clouds to be even faster than that caused by dynamical friction. There is now direct evidence that the centre of the Milky Way contains a stellar bar, a system of stars aggregated in an elongated pattern that rotates about the centre as a solid body. Simulations of gas flow in a barred self-gravitating system suggest that the gas in the bar will fall almost radially towards the centre, shortening by a large margin the amount of time it would take for the gas to get there. However, a detailed calculation for the Milky Way cannot be done until infrared observations of the bar by the Cosmic Background Explorer (COBE) satellite can be interpreted to give a reliable determination of the precise mass distribution of the bar. So it is yet to be determined whether dynamical friction or a barred distribution of stars plays a more important role in the development of the starburst. The issue is an important one because in the latter case, a disruption of the clouds as a result of star formation prior to reaching the nucleus should be relatively less important than if dynamical friction is the primary mechanism by which angular momentum is lost.

Many details are still to be worked out. We do not yet know what physical conditions are necessary for the molecular gas to form the vast numbers of O stars in a starburst. Furthermore, given the reservoir of molecular material at the Galactic Centre, it is not yet possible to predict whether the Milky Way will be either a macho or wimpy version of a starburst. But the general conclusions of the Stark *et al.* paper seem rather robust. Further work should help clarify whether the Milky Way and normal disk galaxies in general all have the capability of becoming starbursts, implying that the starburst phenomenon is an episodic one, or whether only certain galaxies have the right stuff to become starburst. And what's all that molecular stuff doing in the centre of the Milky Way anyway? □

Leo Blitz is in the Astronomy Program, University of Maryland, College Park, Maryland 20742, USA.

TRANSPLANTATION of organs and tissues between individuals is now a routine treatment for a variety of diseases. A drawback of these therapies is immunological rejection of the graft, a response stimulated by antigenic differences between the class I and II HLA molecules of the donor and the recipient. In practice the gravity of the problem varies with the tissue transplanted. Solid organs — kidney, heart and liver — mismatched at several HLA genes are effectively transplanted under cover of immunosuppressive drugs, anti-T-cell antibodies and lymphoid irradiation, whereas the success of bone marrow transplantation is greatly affected by the quality of the HLA match. The extreme sensitivity of bone marrow transplantation to HLA incompatibility is dramatically demonstrated by a case described by Fleischhauer and colleagues in *The New England Journal of Medicine*¹, in which rejection was associated with a single amino-acid substitution in a class I HLA molecule.

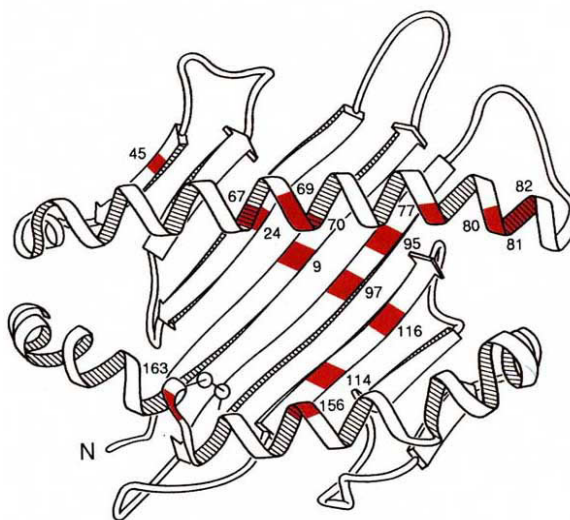
The transplant recipient, who was suffering from chronic myelogenous leukaemia, and the unrelated bone-marrow donor were judged by tissue typing to be HLA identical. This was not the case, however, for electrophoretic comparison of their class I HLA proteins showed the patient to express a more acidic HLA-B44 than the donor. Furthermore, nucleotide sequence comparisons showed the only difference between the two molecules to be at position 156, where aspartic acid in the patient was substituted for leucine in the donor (see figure). Evidence

for cell-mediated graft rejection was observed two weeks after transplantation, and *in vitro* cultures started at that time revealed the presence in the patient's blood of cytotoxic T lymphocytes that killed target cells expressing the donor's, but not the patient's, HLA-B44 subtype. Although, as Fleischhauer *et al.* emphasize, "a cause-effect relationship between these cytotoxic T lymphocytes and marrow-graft rejection has not been proved directly" there are reasons, including many precedents from experimental transplantation in laboratory animals, to believe that is the case.

For over 10 years it has been known that serological tissue-typing does not distinguish certain 'subtypes' encoded by closely related class I HLA alleles², an extreme example being the HLA-B27 antigen which embraces the products of at least seven different alleles³. Such subtypes can readily be distinguished by other assays and in the case of HLA-B44 the two subtypes were first described on the basis of electrophoretic differences and their differential recognition *in vitro* by alloreactive T cells⁴. As more precise molecular methods have been applied to the analysis of HLA genes, the magnitude of the discrepancy between serological definition and molecular reality has increased⁵ — even with the most conservative estimation one must now expect the actual number of class I HLA alleles to exceed the currently defined antigens by a factor of two to three.

The similarity in structure of class I subtypes has led some to suggest that their differences will be of little consequence for transplantation and can therefore be ignored. This assumption works when donors and recipients are being matched within families not because subtype differences are inconsequential, but because the task of typing is reduced to tracking the segregation patterns of entire HLA haplotypes. When an HLA-identical family member is not available, however, and the search extends to panels of unrelated donors⁶, the assumption breaks down, as demonstrated by Fleischhauer *et al.*¹.

There is indeed good reason to believe that HLA-A,B subtypes differ in ways that efficiently stimulate alloreactive T cells and provoke graft rejection. With few exceptions the differences between subtypes are at sites that determine the interactions of class I molecules with antigenic peptides or T-cell receptors (see



Ribbon diagram of the $\alpha 1$ and $\alpha 2$ domains of HLA-B showing the positions of residue 156 in the centre of the α helix of the $\alpha 2$ domain. Comparison of the sequences of 35 HLA-B alleles¹¹ shows that the 16 positions with highest amino acid variability, which are shown as shaded blocks in the diagram and include 156, are all at peptide and T cell receptor-interaction residues of the antigen recognition site. The structure is based on that of HLA-A2 (ref.12).

Solvent solution

In an epidemiological study published in *The Veterinary Record* (**128**, 199–203; 1991), J.W. Wilesmith *et al.* trace the onset of the bovine spongiform encephalopathy (BSE) epidemic in British cattle back to a change in the practice of meat and bone meal production. Around 1980, most British 'rendering' plants stopped using hydrocarbon solvents to separate fat from carcasses. This abrupt change fits in with the supposed transmission to cattle, through their feed, of the infective agent of scrapie (a similar disease afflicting sheep) in the winter of 1981–82. The authors point out that the additive effects of heat treatment and solvent action, or the additional steam heating needed to remove the solvents, may explain the agent's inactivation when solvents are used, and that regional variation in rendering practice could underlie the distribution of BSE cases.

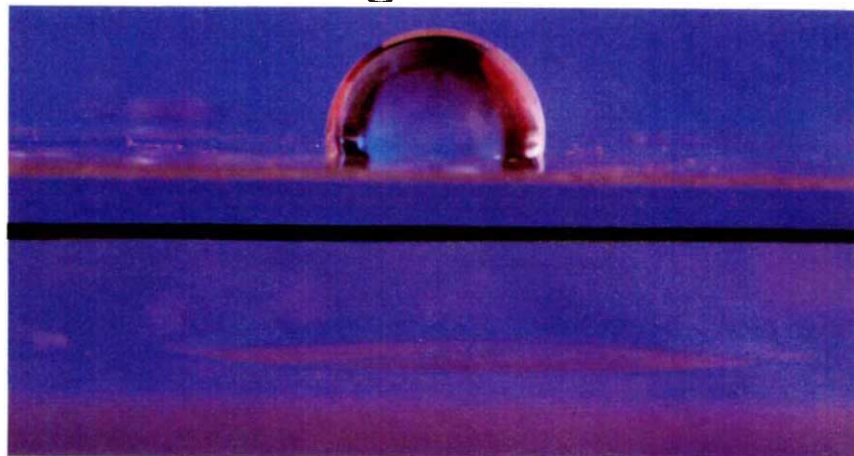
Bated breath

Is the Earth warming up in response to increased atmospheric concentrations of greenhouse gases? The answer, according to T. Karl *et al.* (*Science* **251**, 1058–1061; 1991), is that we won't be able to tell for some time yet. Although the past decade was the warmest on record, it is difficult to distinguish the effects of additional greenhouse gases from natural variations in the climate. Comparing the records of temperature and precipitation for the central United States with model projections for the past 95 years, the authors find no significant indications of warming trends. If the model projections are right, an unambiguous greenhouse signal will emerge only over the next 10–20 years. It will take even longer for any effects of greenhouse warming to show up in levels of precipitation. This leaves policy makers, the authors remark, "in an unenviable position".

Sweet success

Using linkage studies, G.I. Bell and colleagues have mapped the gene responsible for one form of non-insulin-dependent diabetes mellitus (*Proc. natn. Acad. Sci. U.S.A.* **88**, 1484–1488; 1991). First described in the 1970s, maturity-onset diabetes of the young (MODY) is an autosomal-dominant disorder that otherwise closely resembles late-onset non-insulin-dependent diabetes. Using 79 DNA markers, Bell *et al.* excluded about 40 per cent of the human genome, including several candidate genes such as the insulin-responsive glucose transporter on chromosome 17. Close linkage to the MODY locus was finally demonstrated in one large affected family with the adenosine deaminase gene, which maps to the long arm of chromosome 20. Although the location of the MODY gene does not betray its identity, its eventual isolation may provide insight into other forms of diabetes as well.

Fashioning wet-look latex



WHEREAS water on the surface of most artificial plastics and elastomers beads up as shown in the top frame, Isao Noda has developed (page 143) an elastomeric latex that is fully wettable by water (bottom frame; the fully spread droplet is barely visible). The surface of this polystyrene–polybutadiene latex is rendered hydrophilic by adding a diblock copolymer during synthesis, one end of which is water-soluble (polyethylene oxide, PEO) and the other hydrophobic (ensuring solubility in the hydrocarbon latex). For reasons as yet unclear, the amphiphilic copolymer migrates to the surface during polymerization of the latex constituents, there becoming immobilized with the PEO ends exposed to the air.

Block copolymers, used already as emulsifying agents for polymer alloys, are thus able to influence surface as well as bulk properties. The immediate applications (nappies and bandages) may seem prosaic. But might this be a general way to make surfaces wettable only by specific adsorbates? P.B.

figure). Position 156 at which the two HLA-B44 subtypes differ provides a classic example; it is a peptide-interacting residue, the most variable residue of the $\alpha 2$ domain helix, and substitutions at this position have repeatedly been shown to alter T-cell responses. In the course of human evolution class I subtypes have probably been selected for their capacity to bind and present different disease-related peptides to T cells. Substitutions that endow such properties are likely, given the common mechanism of T-cell recognition, to stimulate distinctive alloreactive cytotoxic T-cell responses. In fact that is precisely how subtypes were discovered in the first place². Thus the experience of Fleischhauer *et al.* with HLA-B44 will probably have general application to cases where transplant donors and recipients are mismatched for class I HLA subtypes.

Another feature of the class I HLA subtypes is their distribution in populations of different geographical origins. For example the B*2703 allele has been found only in African and American blacks, the B*2704 and B*2706 alleles only in orientals, and the B*2701 and B*2702 alleles only in caucasians³. These differences, which go undetected by current HLA typing and are likely to provoke allograft rejection, undoubtedly contribute to the difficulty of matching transplant donors and recipients of different racial origins. Such considerations substantiate the view that "The racial differences in ABO blood groups and MHC phenotypes make it more likely that a

candidate for transplantation will receive a well-matched kidney from a member of the same race"⁷.

Through the use of an ever-expanding arsenal of non-specific immunosuppressants, surgeons have been able increasingly to evade the immunogenetic rules of transplantation. As a consequence, the value of HLA matching for solid organs transplanted between unrelated donors is often called into question. For bone marrow transplantation there is in contrast little doubt as to the importance of obtaining the best HLA match, and over the past 20 years the procedure has evolved into the preferred therapy for a widening range of haematopoietic disease^{6,8,9}. This impressive advance has primarily been achieved by the study of bone marrow transplantation between HLA-identical siblings. Now, however,

1. Fleischhauer, K., Kernan, N. A., O'Reilly, R. J., Dupont, B. & Yang, S. Y. *New Engl. J. Med.* **323**, 1818–1822 (1990).
2. Biddison, W. E., Ward, F. E., Shearer, G. M. & Shaw, S. J. *Immun.* **124**, 548–552 (1980).
3. López de Castro, J. A. *Immun. Today* **10**, 239–246 (1989).
4. Hahn, A. W. A., Schendel, D. J., Hansen, P. W. & Ploegh, H. L. *Hum. Immun.* **11**, 69–76 (1984).
5. Bodmer, J. G., Marsh, S. G. E. & Albert, E. *Immun. Today* **11**, 3–10 (1990).
6. Hows, J. M. & Bradley, B. A. *Br. J. Haematol.* **76**, 1–6 (1990).
7. Kasiske, B. L. *et al. New Engl. J. Med.* **324**, 302–307 (1991).
8. Martin, P. J., Hansen, J. A., Storb, R. & Thomas, E. D. *Adv. Immun.* **40**, 379–438 (1987).
9. Santos, G. W. *Clinics Haematol.* **12**, 611–639 (1983).
10. Saiki, R. K., Bugawan, T. L., Horn, G. T., Mullis, K. B. & Erlich, H. A. *Nature* **324**, 163–166 (1986).
11. Bjorkman, P. J. & Parham, P. A. *Rev. Biochem.* **59**, 253–288 (1990).
12. Bjorkman, P. J. *et al. Nature* **329**, 506–512 (1987).

most potential marrow recipients in Europe and North America lack an HLA-identical sibling, and this has led to the establishment of large registries of potential donors from which unrelated HLA-matched individuals can be selected⁶.

The message from Fleischhauer *et al.* is then that HLA typing and assessment of tissue compatibility for bone marrow transplantation will be of the greatest value if all HLA-A, B and C alleles, subtypes included, can be distinguished by routine tissue typing. Moreover the increasing ethnic diversity of North American and European populations

underscores the value of a global approach to defining subtypes. Given the inadequacy of the serological typing methods, meeting these goals will require techniques based upon the nucleotide sequences of the alleles themselves. Such DNA-based typing methods are beginning to be used for class II HLA alleles¹⁰, and the results obtained by Fleischhauer *et al.* must surely catalyse their application to class I. □

Peter Parham is in the Departments of Cell Biology, and Microbiology and Immunology, Stanford University School of Medicine, Stanford, California 94305, USA.

Salvador E. Luria (1912–1991)

S. E. LURIA, who died on 6 February, brought bacteria and their viruses, the bacteriophages, into the forefront of research on the nature of the gene.

Born in Turin, Italy, Luria initially trained to be a physician and chose radiology for his speciality. After military service he moved to Rome to learn more physics in the circle of Enrico Fermi and Edoardo Amaldi. There Franco Rosetti excited him about radiation biology and the formulations by the German physicist Max Delbrück of the gene as a molecule. Equally importantly, he discovered the existence of bacteriophages and began experiments on them in the laboratory of the bacteriologist Geo Rio. Growing official anti-semitism in Italy necessitated his move to Paris in late 1938 where he continued working on phage under the patronage of the physicist Fernand Holweck, doing experiments that aimed to determine the sizes of phages through measurements of the rate at which they were inactivated by ionizing radiations.

The German conquest of France forced him again to flee, this time to New York City which he reached in September 1940 by boat from Lisbon. At the suggestion of Fermi, Luria approached the Rockefeller Foundation and the fellowship they provided allowed him to resume phage work at the College of Physicians and Surgeons of Columbia University. Soon he was in contact with Delbrück, also by then a refugee from totalitarian tyranny. By the summer of 1941, they had begun joint experiments on *Escherichia coli* and its phages, first at Cold Spring Harbor and later at Vanderbilt to which Delbrück had moved to teach physics.

Initially, they focused on the interference phenomenon which permits only one type of phage to multiply in a given bacteria. In the course of their experiments, they used bacteria variants resistant to specific phages. At that time it was an open question whether these resistant cells arose through gene mutations, since traditional wisdom among bacteriologists was that bacteria lacked chromosomes and a form of heredity similar to that of higher organ-

isms. In particular the English physical chemist Cyril Hinshelwood argued against mutational origins of resistant bacteria, believing that they arose through altered chemical equilibria.

It was Luria who provided the key idea to distinguish between the mendelian and lamarckian explanations. What was needed was a way to show that the phage-resistant variants of *E. coli* existed before they came into contact with phage. In February 1943, soon after his move to Indiana University, he realized that the distribution of resistant bacteria in a series of different cultures should settle the question. If the bacteria were made resistant by contact with phage, the number of resistant bacteria would be very similar in all the cultures. But if they arose by mutations, they should be clustered in families, the size of which reflected the time at which the mutational events occurred during the growth of the cultures. Depending upon when the mutations took place, there could be one, two, four, eight, and so on, resistant cells in each culture tube.

As soon as Luria found his first evidence favouring gene mutations he wrote to Delbrück, who then worked out the appropriate mathematical equations. The resulting Luria–Delbrück manuscript, which appeared late in 1943, changed the face of genetics by making bacteria the obvious organisms for research on the nature of the gene. Experiments could be done in a day instead of weeks, and billions of cells could be examined in searching for mutants and determining mutation rates.

The following year Luria went on to show that phages also spontaneously mutate as they multiply, giving rise to variants every bit as stable as those found in bacteria. Two years later, Delbrück and Alfred Hershey independently provided evidence for genetic recombination in phage, while Joshua Lederberg and Edward Tatum demonstrated genetic recombination in *E. coli*. Thus, within a brief three-year span, the existence of chromosomes within both *E. coli* and its phages was established. For these experiments and Hershey's subsequent demonstration that DNA was the

genetic component of phages, Luria, Delbrück and Hershey in 1969 received the Nobel Prize in Medicine and Physiology.

Less appreciated was the key role played by Luria in the 1952 discovery of the host-modification phenomenon in which the host range of a progeny phage can be influenced by the exact strain of bacterium in which it has multiplied. A decade later, Werner Arber went on to show that such behaviour reflects enzymatic attack on unmodified phage DNA and its prevention by methylation.

Luria was an exceptionally talented writer. His scientific papers, textbooks and books for the general public all reflect mastery of his adopted English language. His first popular book, *Life: The Unfinished Experiment* (1973), won the National Book Award, with his autobiography *A Slot Machine, A Broken Test Tube* (1984) lucidly describing his intellectual and humanistic development. I remember him to be a teacher of the first rank. In autumn 1947, after only a few days into his course on viruses, I wanted to do my PhD under his supervision. It was typical of his devotion to his students' future success that he later arranged with John Kendrew for me to go to the Cavendish Laboratory where I was to meet Francis Crick.

Equally important was his skill as an administrator, shown over the 13 years he wisely and compassionately directed MIT's newly formed Cancer Center.

He was a human of passionate political beliefs. From early youth he identified himself with the causes of organized workers as opposed to those of management. As the Vietnam war developed, he became increasingly prominent in the antiwar movement and refused payment of the war-related component of his income taxes. Oppression of any form disturbed him, and there was never any doubt as to where he stood about authoritarian governments or institutions. More recently he worried about society's handling of the human genetic data whose accretion would be greatly accelerated by the human genome project. The possibility of a genetically defined underclass disturbed him, and we disagreed as to whether our societies would find the means to protect the victims of unjust throws of the genetic dice.

Luria knew what science at its best was and strove at all times to maintain high standards in both science and the human behaviour that make it possible. In doing so while young he could offend out-of-date minds or those who were all too self-satisfied with their accomplishments. By the time he died, however, there were few who did not feel better by being in his presence.

JAMES D. WATSON

James D. Watson is at the Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA, and the National Center for Human Genome Research of the National Institutes of Health.

As quick as a flash

Paul Calvert

To get the most out of the large capacity of optical fibre communication networks, one wants to be able to switch the signal on and off as fast as possible to transmit the maximum number of noughts and ones. Liquid crystals take several milliseconds to switch, and so are too slow. Lithium niobate is now used and can switch in fractions of a nanosecond. But to get to the desirable switching times of a picosecond or less, one has to rely on the rapid movement of electrons in molecules. Conjugated aromatic molecules offer the most promise for large changes of refractive index, induced by a redistribution of electrons in an electrical or optical field. However, nothing wholly suitable has yet been found. Nothing daunted, C. S. Winter *et al.*, from the British Telecom Research Laboratories, now suggest¹ that nickel dithiolene shows great promise.

As described in a previous News and Views², one possible optical switch is a Mach-Zehnder interferometer in which a light signal is split into two: one side travels through an optically active material whose refractive index is modified by an electric field. The increased index slows the light in this arm until it out of phase with the signal in the other arm when they recombine. With the field on, the two arms subtract to give zero; with the field off they add.

Stegeman *et al.*³ have proposed figures of merit for nonlinear optical materials to balance the efficiency of refractive index change against the loss of signal due to scattering or absorption. Unfortunately the best available material on this basis is silica which has a low nonlinear coefficient but extremely low optical absorption and scattering. A device made from silica would have to be enormously long to get the half-wavelength retardation needed to switch the signal off.

Materials can be made with very large nonlinear coefficients (χ_3) if they have an absorption peak at the wavelength of the test laser, but the figure of merit is low because of the high absorption. A better approach might be to move the peak absorption wavelength away from the laser wavelength (usually Nd-YAG light at 1.064 μm) until scattering of the light by nonuniformities in the film becomes the dominant loss mechanism. Winter *et al.*¹ suggest a compromise of putting the absorption peak at slightly shorter wavelengths than the laser and have tested two nickel dithiolate compounds absorbing at around 750 nm. The combination of a big nonlinearity and low absorption gives a figure of merit, W , of about 5. This figure, W , is the maximum refractive index change, at light intensities high enough to start damaging the material, divided by the absorption coefficient and the wavelength. It should be at least 3 and ideally more than 12 for good switching.

The materials most studied to date are the polydiacetylenes. Measurements on the *p*-toluenesulphonate substituted diacetylene produced a W of 0.15, away from the absorption³. Polydiacetylene films can be prepared by direct polymerization of a film of monomer but this results in high levels of scattering. Changing to a soluble polydiacetylene allows the film to be spin-coated onto a surface from solution, with much more uniform films and less scattering. Poly-4BCMU (another diacetylene derivative) can be put down in this way with a W that has been quoted⁴ as 1.5. Unfortunately there is a catch. As the scattering is reduced two-photon absorption becomes a problem. This arises because the very high laser intensities allow absorption corresponding to twice the photon energy of the laser (532 nm wavelength). As this is in an absorption band of the red polydiacetylene, two-photon absorption is quite strong. Original estimates suggested that W could reach as high as 100, but this now looks unlikely.

The nickel dithiolates studied by Winter *et al.*¹ show low absorption at 532 nm and two-photon absorption does not seem to be significant. A further improvement in their efficiency might come from tuning the structure to shift the absorption peak up or down to optimize the balance of absorption and nonlinearity. It will also be necessary to build them into a polymer structure so that good clear films can be readily cast from solution. Scattering is prohibitively high for the films of crystals that can be made now. This also means that solution measurements on dithiolates are not really a fair comparison with actual waveguides made of diacetylenes because real waveguides are likely to have much more scattering, especially with early experimental samples.

Given all this it is not obvious that a W of 5 from a solution is going to be better than the diacetylene value of 1.5 measured on a real film. It is also likely that other undesirable effects, including degradation, will manifest themselves as we get closer to actual devices. It is clear that we are currently tuning molecular structures to optimize a fairly simple combination of spectral properties and we ought to have more predictive ability than we seem to have. What have the theoretical chemists been doing all these years? □

Paul Calvert is at the Arizona Materials Laboratories, University of Arizona, Tucson, Arizona 85712, USA.

1. Winter, C. S., Hill, C. A. S. & Underhill, A. E. *Appl. Phys. Lett.* **58**, 107–109 (1991).
2. Calvert, P. *Nature* **337**, 408 (1989).
3. Stegeman, G. I., Zanoni, R. & Seaton, C. T. *Mat. Res. Soc. Symp.* **109**, 53–64 (1988).
4. Townsend, P. D., Jackel, J. L., Baker, G. L., Shelburne, J. A. & Etemad, S. *Appl. Phys. Lett.* **55**, 1829–1831 (1989).

Food on the hop

Now that biodegradable plastics are at last becoming available, Daedalus is widening the search. He is trying to produce a good biodegradable rubber. DREADCO's chemists point out that like rubber, many proteins exist in a folded-chain configuration, and some are highly elastic — like water-plasticized gelatin (table jelly) and elastin (the energy-storing component of the leg tendons of jumping animals like fleas and Kangaroos). The shape taken up by a given protein molecule is hard to predict, so rational design and synthesis of an elastic protein would be forbiddingly difficult. Instead, the DREADCO team are partially hydrolysing elastin, gelatin and other promising natural proteins, and recombining the fragments into novel 'artificial' (rather than fully synthetic) materials. Already they have produced a protein rubber so biodegradable that it is actually edible.

It seems a pity to waste such an unusual product on tyres, insulation and erasers, so DREADCO's marketers plan to launch it as a foodstuff. They are reluctant to label it as a mere imitation meat. 'Substitute' or 'ersatz' products seldom acquire much dignity or sales; besides, the broiler industry already has a commanding lead in the production of rubber chicken. So the hunt is on for some uniquely valuable application for elastic food. Suggestions made to date include: pre-stressed 'novelty nibbles' that snap and ricochet beguilingly around inside the mouth when bitten; an all-edible elastic-powered clockwork mouse to give domestic cats the thrill of the chase without its cruelty; edible balloons for children's parties and edible rainwear for fetishists; boil-in-the-bag products that let you eat the bag as well; and (inspired by the pop-up toaster) a pre-stressed steak and a jumping sausage with meltable internal restraints, which leap dramatically out of the frying pan when they have reached the desired cooking temperature. But the main customers for specialized, outlandish, expensive foods are, of course, slimmers.

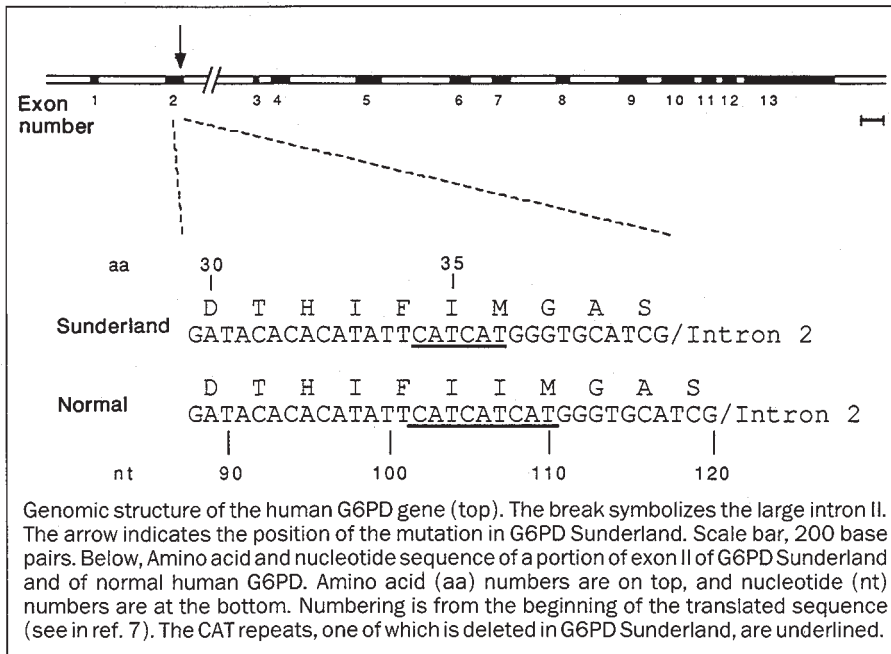
A number of slimming foods are designed to absorb water in the stomach, thus swelling up and giving the impression of fullness. Daedalus is devising an edible foam rubber which looks a bit like brown rice, but with grains that swell immensely once they have been swallowed. The slowly digestible foam core of each piece is compressed within a rapidly digested protein-rubber coating. In the stomach, the coating digests and ruptures, the foam expands wildly, and the consumer feels misleadingly full. Until the expanded pieces have been digested, they cannot even proceed into the narrow entrance to the small intestine. As long as they fill and block his stomach, the consumer continues to feel completely bloated. It may be hours before he can face another helping. David Jones

Deficiency in red blood cells

SIR — Glucose 6-phosphate dehydrogenase (G6PD) is a ubiquitous enzyme encoded by a gene that is highly polymorphic in humans¹. Its deficiency is common in populations that have been subjected to selection by *Plasmodium falciparum* malaria². Common variants entail the risk of acute haemolytic anaemia, but clinical expression of these variants is generally mild. By contrast, some

Kanno *et al.* This part of exon II would not be expressed in red cells if the claim of Kanno *et al.* had been correct. Our observation demonstrates that a mutation in this X-linked amino-terminal region of G6PD causes deficiency in red cells.

This new type of G6PD, which we designate G6PD Sunderland, is the first variant found to result from a deletion rather than a



individuals have rare variants associated with much more severe haemolysis³. The gene encoding G6PD has been mapped on the X-chromosome in mammals, specifically on Xq28 in humans⁴.

Kanno *et al.* claimed⁵ that G6PD in red cells consisted of an amino-terminal region encoded by a gene on chromosome 6, and of a carboxy-terminal region encoded by the X-linked gene. In a News and Views article written at the time⁶, L.L. commented that, although several arguments could be made against the likelihood of such a fusion protein, formal genetic evidence for X-linkage of the amino-terminal region of G6PD was not available, because none of the mutations reported at the time⁷ mapped to that region. The claim by Kanno *et al.* has since been refuted by molecular analysis^{8,9} and attributed to a purification artefact¹⁰. Nevertheless, mutations in the amino-terminal region have still failed to turn up¹¹.

Using a polymerase-chain-reaction based technique¹², we have now determined the nucleotide sequence of the entire coding region of the G6PD gene from a person with severe red cell G6PD deficiency and chronic haemolytic anaemia. The only abnormality we found was a three base-pair deletion in exon II (see figure), which predicts the loss of one of two adjacent isoleucine residues (amino acid 35 or 36), just upstream of the methionine residue called 'junctional' by

base change. The deletion is within a three-fold CAT repeat, and has presumably arisen through misalignment at meiosis, with conservation of the reading frame.

DONALD MACDONALD
MARGARET TOWN
PHILIP MASON
TOM VULLIAMY
LUCIO LUZZATTO

Department of Haematology,
Royal Postgraduate Medical
School,
Hammersmith Hospital,
London W12 0NN, UK

D.K. GOFF

Royal Infirmary,
Sunderland SR2 7JE, UK

1. Luzzatto, L. & Battistuzzi, G. *Adv. hum. Genet.* **14**, 217–329 (1985).
2. Luzzatto, L. *Blood* **54**, 961–976 (1979).
3. Luzzatto, L. & Mehta, A. in *The Metabolic Basis of Inherited Disease* (eds Scriver, C. R., Beaudet, A. L., Sly, W. S. & Valle, D.) 2237–2265 (McGraw-Hill, New York, 1989).
4. Pal, G. S., Sprengle, J. A., Do, T. T., Maren, C. E. & Migeon, B. R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2810–2813 (1980).
5. Kanno, H., Huang, L.-Y., Kan, Y. W. & Yoshida, A. *Cell* **58**, 595–606 (1989).
6. Luzzatto, L. *Nature* **341**, 286–287 (1989).
7. Vulliamy, T. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 5171–5175 (1988).
8. Mason, P. J., Bautista, J., Vulliamy, T. J., Turner, N. & Luzzatto, L. *Cell* **62**, 9–10 (1990).
9. Beutler, E., Gelbart, T. & Kuhl, W. *Cell* **62**, 7–8 (1990).
10. Yoshida, A. & Kan, Y. W. *Cell* **62**, 11–12 (1990).
11. Beutler, E. *Semin. Hemat.* **27**, 137–164 (1990).
12. Pogg, V., Town, M., Foulkes, N. S. & Luzzatto, L. *Biochem. J.* **271**, 157–160 (1990).

Sea-level constraints

SIR — The authors of three important papers^{1–3} used Barbados coral cores of latest Pleistocene and Holocene ages to provide important constraints on sea-level fluctuations and to improve the radiocarbon timescale in this interval by calibrating it against the uranium–thorium timescale. The results of Bard *et al.*³, for example, indicate that the last deglaciation started about 3,000 years earlier when referenced to the U–Th timescale than when referenced to the radiocarbon timescale. These authors suggest that this difference may explain the previously noted discrepancies between observations of sea-level change and models of glacial rebound in which ice sheets are considered only over Laurentia and Fennoscandia and that it is not necessary to introduce meltwater into the oceans from other ice sheets^{4–6}.

This conclusion is incorrect. Time enters into the formulation in three ways: through the history of the changes in the ice sheets; through the observations of sea-level change and through the Earth's rheology. Provided the timescales are linear relative to each other and provided that the same timescale is used throughout, the use of either the U–Th or the ¹⁴C timescale is immaterial. The results of Fairbanks¹ indicate that the relation between the ¹⁴C and U–Th timescales is effectively linear for the past 20,000 years and that any departures from this linearity are small when compared with the uncertainties associated with many of the radiocarbon ages used to constrain ice models and sea-level change.

The ice models are almost wholly constrained by radiocarbon ages of debris left by the retreating ice front. In some instances in Fennoscandia, varve chronologies have been used; these should be reduced to the radiocarbon timescale. (This does not appear to have been done but there are only few varve dates so they are unlikely to distort the global ice-sheet models.) Observations of sea level have sometimes been related to varve or U–Th timescales, but the bulk of reported data refers to the conventional radiocarbon timescale. In the rebound modelling, the Earth's viscous time constraint(s) are usually considered as unknowns so that the viscosities will be either in Pa (¹⁴C)s or Pa (U–Th)s, depending on which timescale is used for the other two inputs. (The distinction is not so insignificant if the two timescales differ by about 3,000 years after 20,000 years.)

If the U–Th timescale for the Barbados corals is used to constrain the glacial-hydroisostatic rebound models then the ice models must be transformed to this timescale as well, with the result that the relation between observation and model prediction will be unchanged unless the relation between the two timescales is significantly nonlinear. One cannot usefully compare the U–Th

timescale-controlled sea-level data with rebound predictions based on ice models linked to the ^{14}C scale. Yet this is what Bard *et al.* do.

A second difficulty with the conclusion of Bard *et al.* is the sea level during the last glacial maximum. Around 18,000–20,000 years ago, sea levels are generally believed to have been at about 130 m below their present level (although this depth will be regionally variable) and the Barbados coral data support this: the minimum depths at this time of the corals are about 120 m below present sea level. Ice models for the Laurentide and Fennoscandia that produce rebound at the centres of these ice sheets that are consistent with these observations of sea-level change within and near the ice-sheet margins contain only enough ice to raise sea level by 60–80 m (refs 5,7). Increasing the volumes in these ice sheets results only in implausible predictions of rebound at these sites unless the melting was initially very rapid. The Barbados sea-level curve from 18,000 to 6,000 years before present, however, argues against this, so it has been suggested that

this missing melt water has to come from other areas such as the Barents–Kara ice sheet in Antarctica, or possibly from eastern Siberia^{4–6}. Rather than negate the previously drawn conclusions about the importance of these centres of deglaciation, the Barbados results reinforce them.

KURT LAMBECK

Research School of Earth Sciences,
Australian National University,
Canberra, ACT 2601,
Australia

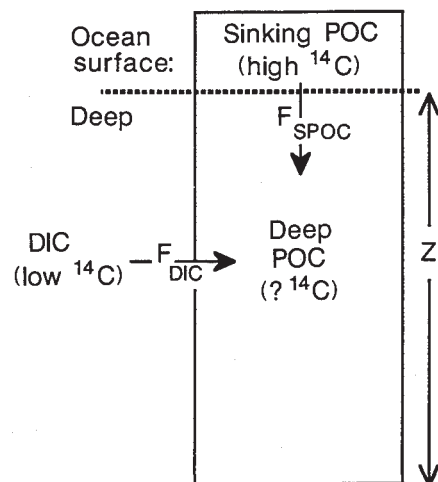
MASAO NAKADA

Department of Geology,
Faculty of Science,
Kumamoto University,
Kumamoto 860, Japan

1. Fairbanks, R. G. *Nature* **342**, 637–642 (1989).
2. Bard, E., Hamelin, B., Fairbanks, R. G. & Zinder, A. *Nature* **345**, 405–410 (1990).
3. Bard, E., Hamelin, B. & Fairbanks, R. G. *Nature* **346**, 456–458 (1990).
4. Nakiboglu, S. M., Lambeck, K. & Aharon, P. *Tectonophysics* **91**, 335–358 (1983).
5. Nakada, M. & Lambeck, K. *Nature* **333**, 36–40 (1988).
6. Peltier, W. R. *Quat. Res.* **29**, 93–112 (1988).
7. Peltier, W. R. & Andrews, J. T. *Geophys. J. R. astr. Soc.* **46**, 605–646 (1976).

Another recipe for bomb ^{14}C dilution

SIR — As recently pointed out by Druffel and Williams¹ and Toggweiler², the relatively low amount of atom-bomb-derived ^{14}C in deep-ocean particulate organic carbon (POC) has significant implications for the way carbon is cycled in the ocean's interior. Druffel and Williams conclude that the low $\Delta^{14}\text{C}$ found



Model of deep-ocean POC formation from dissolved inorganic carbon (DIC) and sinking POC sources. Within the deep ocean, sinking and suspended POC are not differentiated. F_{SPOC} , areal flux rate of sinking POC through the surface/deep ocean boundary; F_{DIC} , areal flux rate (uptake) of deep DIC into POC.

in POC suspended in the deep ocean is probably the result of dilution by incorporation of ^{14}C -depleted dissolved organic carbon (DOC) into deep POC. As well as DOC→POC transformations, I wish to point out that dissolved inorganic carbon (DIC) is another large carbon source for such

dilution, and that the conversion rate of DIC to POC in the deep sea is quantitatively, and therefore isotopically, important.

Let us assume that the only source of non-bomb carbon in the deep ocean is DIC, and that deep POC must therefore originate from some combination of this source and ^{14}C -rich sinking POC (see figure). If we know the fluxes and $\Delta^{14}\text{C}$ of these sources, the $\Delta^{14}\text{C}$ in POC resident in the deep ocean can be calculated. Although it has been assumed¹ that uptake of DIC into POC is negligible in the deep sea, Sorokin reported³ uptake rates ranging from 0.5 to 3 $\mu\text{g C m}^{-3} \text{ d}^{-1}$ at depths greater than 1 km in the central Pacific Ocean. Such rates are indeed minuscule when compared to those in surface waters, but they become significant when integrated over depth and compared with the sinking POC flux into the deep ocean. Using a mean deep-ocean DIC uptake rate of 2 $\mu\text{g C m}^{-3} \text{ d}^{-1}$, the estimated areal flux rate of DIC into POC (10 $\text{mg C m}^{-2} \text{ d}^{-1}$) equals the sinking POC flux into deep water in the north central Pacific Ocean (see table). This yields an estimated deep POC $\Delta^{14}\text{C}$ of -52‰ (see table), even more depleted in bomb ^{14}C than actually observed¹ ($> 43\text{‰}$), indicating the potential importance of the DIC→POC pathway in effecting ^{14}C dilution in the deep sea.

In the Santa Monica Basin, a greater flux (30 $\text{mg C m}^{-2} \text{ d}^{-1}$) of POC sinking into a shallower aphotic zone could be viewed as less advantageous for dilution of bomb ^{14}C in deep POC. But aphotic DIC→POC conversion in such basin environments ranges from < 1 to 24 $\text{mg C m}^{-3} \text{ d}^{-1}$ (refs 4–7), at least several orders of magnitude greater than those observed³ in the open ocean. Assuming a conservative mean rate estimate for this site of 1 $\text{mg C m}^{-3} \text{ d}^{-1}$, depth-integrated uptake of deep DIC is calculated to be 800 mg C

$\text{m}^{-2} \text{ d}^{-1}$ or more than 20 times the flux entering aphotic waters via sinking POC. This again yields an estimated deep POC $\Delta^{14}\text{C}$ value that is much lower than that of POC sinking from surface waters (see table).

A further example of the magnitude of such a process is provided by Tuttle and Jannasch⁵, who reported that DIC uptake in waters above the Cariaco trench is at least 17–58 % as great as photosynthetic inorganic carbon assimilation in overlying surface waters. The above observations again lead to the conclusion that bomb ^{14}C in

FLUXES AND $\Delta^{14}\text{C}$ RELEVANT TO DEEP POC FORMATION

Site	Z (m)	Fluxes ($\text{mg C m}^{-2} \text{ d}^{-1}$)		$\Delta^{14}\text{C}$ (‰)		DPOC
		SPOC	DIC	SPOC	DIC	
NCP	4,760	10	10	+136	−240	−52
SMB	800	30	800	+86	−50	−45
Ref.	1	8,9	3–7	1	1	(This report)

SPOC, sinking POC at surface/deep ocean boundary; DPOC, deep POC (both sinking and suspended); NCP, north central Pacific Ocean site²; SMB, Santa Monica Basin site². The DIC flux (F_{DIC}) is derived by multiplying the mean per-volume rate estimates (2 $\mu\text{g C m}^{-3} \text{ d}^{-1}$ (ref. 3) and 1 $\text{mg C m}^{-3} \text{ d}^{-1}$ (refs 4–7) for NCP and SMB, respectively) times the respective deep-ocean water column depth, Z. The surface/deep ocean boundary is arbitrarily chosen to be at a depth of 1,000 m at the NCP site and at 100 m at the SMB site. $\text{DPOC } \Delta^{14}\text{C} = ((F_{\text{SPOC}}/F_{\text{DIC}}) \times \text{SPOC } \Delta^{14}\text{C}) + ((F_{\text{DIC}}/F_{\text{DIC}}) \times \text{DIC } \Delta^{14}\text{C})$ where $F_{\text{DIC}} = F_{\text{SPOC}} + F_{\text{DIC}}$ (ref. 1) in units of parts per thousand (‰).

resident deep-ocean POC can be substantially diluted by DIC uptake, independent of DOC incorporation or other mechanisms¹.

Thus, as previously noted¹⁰, DIC uptake should not be ignored in interpreting the paucity of ^{14}C in some deep-sea organisms and POC pools. It would therefore be premature to ascribe such ^{14}C depletion to DOC incorporation alone. Clearly, further rate and isotope measurements are needed better to characterize this important piece of the global carbon cycle.

GREG H. RAU

Institute of Marine Sciences,
University of California,
Santa Cruz, California 95064 USA

1. Druffel, E. R. M. & Williams, P. M. *Nature* **347**, 172 (1990).
2. Toggweiler, J. R. *Nature* **347**, 122 (1990).
3. Sorokin, Yu. I. *Oceanology* **11**, 85 (1971).
4. Seki, H. & Zobell, C. E. *J. Oceanogr. Soc., Japan* **23**, 182 (1967).
5. Tuttle, J. H. & Jannasch, H. W. *Limnol. Oceanogr.* **24**, 746 (1979).
6. Seki, H. *et al. Arch. Hydrobiol.* **100**, 73 (1984).
7. Juniper, S. K. & Brinkhurst, R. O. *Mar. Ecol. Prog. Ser.* **33**, 41 (1986).
8. Martin, J. H. *et al. Deep-Sea Res.* **34**, 267 (1987).
9. Nelson, J. R. *et al. Cont. Shelf Res.* **7**, 307 (1987).
10. Rau, G. H., Karl, D. M. & Carney, R. S. *Deep-Sea Res.* **33**, 349 (1986).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. They need not arise out of anything published in *Nature*. In any case, priority will be given to letters of fewer than 500 words and five references. □

Disease of war

Alastair Hay

Preventing a Biological Arms Race. By Susan Wright. MIT: 1990. Pp.446. £22.50, \$25.

If an image were required to sum up what biological warfare is about, it could be the remnants of the bombed baby-milk factory on the outskirts of Baghdad. The United States claims that it was a biological weapons plant, but Iraq says it was nothing other than a factory formulating powdered milk into infant feeds. Who do you believe?

A journalist on the *Washington Post* diligently tracked down the plant manufacturer (in New Zealand) which maintains that it was indeed built to process powdered milk. Apparently the factory was closed for a time and then restarted by French engineers, again to handle powdered milk. But, says the United States, it was after August 1990, and after the French had left, that the factory was used for nefarious purposes. And why, the Americans ask, was it in camouflage livery? Iraq says the camouflage can be explained simply by the availability of building material when the plant was constructed.

There is a sense of unreality about this. In fact much of the ruminating about biological warfare is unreal. So much information is shrouded in a veil of secrecy. At a time like this, anything that serves as a reminder that it is not just make-believe, or not all make-believe, is welcome. *Preventing a Biological Arms Race* is just such a reminder. Edited by Susan Wright, the book is a systematic collection of essays written by participants in the biological warfare debate. Few stones are left unturned in chapters that range over the origins of the US Biological Warfare Program (the United States destroyed its biological weapons between 1970 and 1973), the potential of genetic engineering to modify existing diseases for use in munitions, the role of international law and, of course, allegations about biological warfare.

Fortunately, for the moment, we are still only talking about allegations. But when the United States accuses the Soviet Union of violating the two international treaties that govern conduct in this area, the 1925 General Protocol and the 1972 Biological Weapons Convention, both of which outlaw biological warfare, the allegations must be taken seriously.

According to the United States, the first foul occurred in the Soviet city of Sverdlovsk in 1979. The problem was anthrax, caused, says the United States, by the release of airborne spores following an explosion at a biological weapons facility. Release of the spores was said to have caused pneumonia in some people and, therefore, the most likely candidate was pulmonary anthrax, an agent considered by the United States at one time as a suitable biological weapon. The Soviet response acknowledged that an outbreak of

anthrax had occurred, but claimed that it was gastric in origin, and caused by the consumption of contaminated meat. Anthrax is still endemic in parts of the Soviet Union.

Support for the more prosaic Soviet version of events is provided in an interview given to the *New York Times* in 1981 by Donald Ellis, then professor of physics and chemistry at Chicago's North Western University. Ellis was in Sverdlovsk at the time of the alleged incident, and said that he regularly passed unhindered within a few blocks of the site on his way to and from

Biological weapons plant or baby-milk factory — who do we believe?

work. Ellis felt that if there had been a release of anthrax spores, the Soviet authorities would have taken steps to protect him. Rumour has it that Ellis was not the only visiting foreign scientist; half a dozen others were also in the city at the time and were apparently allowed to move freely about its precincts.

Regrettably, the Soviet authorities neither invited senior US scientists to visit the site of the accident nor did they offer them medical records to examine. Prudence might have suggested that an invitation would have been wise, as allegations have a tendency to stick unless refuted. Although there is no requirement under the existing treaties for countries to allow inspections of sites that are the subject of dispute, the Soviet reluctance to have outsiders take a look has only fueled speculation that it had something to hide.

In view of the seriousness of the Sverdlovsk incident, it would have been extremely valuable to have had a Soviet contributor to this book addressing some of the issues raised in its chapters. Had such a contributor been found, he/she would rightly have poured scorn on the second of the American allegations, that is that the Soviet Union had

encouraged the use of fungal toxins as chemical/biological warfare agents in Southeast Asia between 1978 and 1983. These incidents, known in the trade as 'yellow rain' following the description — by a Hmong soldier from Laos — of an alleged spray attack, have come to resemble a farce. We now know, improbable though it seems, that what was thought to be a sophisticated delivery system for toxins using deproteinized pollen and uric acid, was nothing other than the excreta of Asian honey bees.

Yellow rain has now been thoroughly debunked. Scientists and others in many countries have assisted in the systematic dismantling of the US case. The most effective sceptic was Matthew Meselson, professor of biochemistry and molecular biology at Harvard. In a chapter that explains the allegations and the evidence for and against, Meselson and two colleagues point out how the Reagan administration locked itself into a "political position from which it has yet to extricate itself". The administration had only itself to blame. By failing to corroborate testimonies of witnesses, ignoring pleas by Department of Defense scientists for more time to confirm findings, and bypassing its own scientific advisers, the administration broke all bureaucratic rules by going public on the flimsiest evidence.

Internal rules partly govern the way governments behave, but international rules are also needed to keep them in check. Given its use of chemical weapons against Iranian soldiers and Kurdish civilians, Iraq clearly needs the constraints of international opinion. But from this distance it seems improbable that Iraq would house a biological weapons plant in a corrugated iron building. In view of the serious risk of local contamination it is even more questionable that Iraq would then allow journalists to trample over the bombed site to witness torn sacks of powdered milk. A far more likely site, if indeed there is one, would be the pharmaceutical plant said to be two miles down the road.

Whether or not Iraq has biological weapons remains a subject of speculation. Sadly, what is no longer just speculation is the growing body of evidence — amply documented in this book — that biological warfare is being viewed again by some sections of the military as something that may be of value. The counterweight to this is the evidence that most governments respect, and wish to strengthen, the Biological Weapons Convention. The chance to do this will occur later this year when the convention undergoes its third review (the last was five years ago). The contributors to *Preventing a Biological Arms Race* have helped to focus attention on the additional measures that are needed to support the convention, and to thwart the military interest in biological warfare. For thwarted it must be. □

Alastair Hay is in the Department of Chemical Pathology and Immunology, Old Medical School, Leeds LS2 9JL, UK.

Mostly about aluminium

Robert W. Cahn

R & D for Industry — A Century of Technical Innovation at Alcoa. By Margaret B. W. Graham and Bettye H. Pruitt. Cambridge University Press: 1991. Pp.645. \$49.50, £50.

SOME years ago, the *New Yorker* carried one of its 'cocktail party' cartoons: a group of stolid businessmen is conversing, and one declares: "I've learnt a lot in 62 years, but unfortunately most of it is about aluminum." Drs Graham and Pruitt have written a massive book that on the face of it is also mostly about alumin(i)um — but it is also about much more. The authors (one of whom is professor of operations management at Boston University) are associated with the Winthrop Group, a company which specializes in the history of business and technology. The book, in the guise of an exhaustive history of the research laboratory of the Aluminum Company of America, is really about the permanent, crucial and insoluble problem of the proper relationship between commercial management and research management in a large industrial company. As the authors point out, aluminium as an industrial material, and corporate industrial research and development are both about a century old; their crises and triumphs developed in synchrony.

Although the German chemical industry already had large research laboratories in the 1880s, it is generally accepted that the first true corporate industrial R&D laboratory in America was set up in a Schenectady barn in 1900 by Willis R. Whitney for the General Electric Company (see the splendid biography *Willis R. Whitney, General Electric and the Origins of US Industrial Research* by George Wise, Columbia University Press, 1985); before that, there had only been troubleshooting laboratories attached to individual works. After 1900, the idea of

segregating corporate researchers in a separate location, sheltered to some degree from day-to-day operating crises, rapidly gained ground. Du Pont already had its Experimental Station in 1902, and its history has been exhaustively analysed in another excellent book recently published by the Cambridge University Press (David A. Hounshell and John K. Smith, *Science and Corporate Strategy, Du Pont R&D, 1902–1980*, 1988). Most of the points made about that very different company find their echoes in the new book about Alcoa.

Alcoa's business was built on the patent granted in 1889 to Charles Martin Hall, the co-inventor (with Frenchman Paul Héroult) of the electrolytic process still used today for smelting aluminium. Hall became a technical director of the fledgling company, a position which he used to run a small group of subordinates within a factory, under his close direction. He defended his personal authority with tigerish determination when he refused, in 1909, to permit a corporate research laboratory to be established, a determination which he maintained even when he was dying of leukaemia some years later. In consequence, the laboratory was established only in 1919, and the company suffered greatly from this late decision. The battle of Hall versus the rest of the board of directors was a precursor of many later battles about the size, ethos and degree of independence of the company's successive R&D laboratories, battles which swung first one way and then another, culminating in the construction of the Alcoa Technical Center, where research (development moved in earlier) began in 1972.

There is space here only to list some of the perennial issues discussed in the light of the company's history: project control (in R&D) by a system of standing committees set up by commercial management versus a role for long-range research under the control of a director of research; protection of company interests by patenting versus secrecy without patents; whether it was safe to accept Government money for research (for years it was thought unsafe, a direct consequence of the Justice Department's initiation of anti-trust action in 1937, an action which dragged on for years and obliged Alcoa to hand over a number of patents). Other issues included whether researchers could safely be encouraged to participate fully in the professional life of their various scientific specialities versus a policy of purdah, to protect company secrets; developing detailed processing technology for novel uses of the company's products versus leaving that to customers (this was a very major issue over the introduction of aluminium alloys for beverage containers); what resources it was proper to devote to the effort to replace the Hall-Héroult process by a radically different way of making aluminium (this expensive and sustained effort came to nothing, but during its time powerfully buttressed the reliance on secrecy with all its

attendant problems).

The later part of the book recounts the decision in the 1980s to recast research strategy in a very radical way, partly as a result of five years' pressure by an outsider brought in to advise on R&D policy. Long-range research was instituted to give the company completely new options to diversify, in due course, from aluminium, greatly increasing total R&D expenditure and also the number of patents taken out; the number of staff with doctorates was doubled, which of necessity led to increased involvement of staff in the activities of their professional bodies, and obliged Alcoa to institute an in-house system of awards to distinguished researchers. In spite of the success of the new R&D strategy, there was further boardroom anxiety about the cost, leading in 1987 to the replacement, by an outsider, of the company chairman who had masterminded the strategy. The creative tension between the commercial ethos and the research ethos, with right on both sides, can never be finally resolved; the current problems at the Philips Research Laboratories show just how difficult this tension is.

This is a profound and stimulating book, always analytical, never hagiographic, rarely censorious. It is a distinguished part of the new wave of industrial history, and well worth the attention of any applied scientist, irrespective of his speciality. □

Robert W. Cahn is in the Department of Materials Science and Metallurgy, Cambridge University, Pembroke Street, Cambridge CB2 3QZ, UK.

Multiple scanning probes

Graham Dunn

Confocal Microscopy. Edited by T. Wilson. Academic Press: 1990. Pp.426. £33, \$71.

IN a charming memoir recounting his invention of the confocal scanning microscope in 1955, Marvin Minsky recalls that it was fortunate that his brother-in-law was a patent attorney, otherwise the instrument might never have been documented at all. It was fortunate indeed for biologists because, in combination with lasers, confocal scanning microscopy has now become indispensable for exploring the three-dimensional structure and dynamics of living systems. In the words of Shinga Inoué: "Seldom has the introduction of a new instrument generated as instant an excitement amongst biologists..." Even so, the full potential of confocal scanning microscopy is still far from realized and this latest book on the subject contains a fascinating insight into possible future developments of this very versatile technique.

Confocal Microscopy starts with a heavy chunk of theory — indigestible for many, but

New in paperback

The Origins of our Universe by Malcolm S. Longair is based on the authors 1990-91 Royal Institution Christmas Lectures. The book is published by Cambridge University Press; price £7.95, \$12.95 (also available in hardback; price £20, \$34.50).

Fields, Strings and Quantum Gravity edited by Hanying Guo, Zhaoming Qiu and Henry Tye is published by Gordon and Breach and examines the present status of string theory, quantum gravity and conventional field theory in theoretical physics. Price is \$105.

Chapman and Hall recently published *Monitoring for Conservation and Ecology* at a price of £14.95, \$35. The editor, F. B. Goldsmith brings together a team of ecologists and conservationists to provide a useful summary of ecological monitoring.

invaluable to readers interested in developing a confocal microscope or getting the most out of an off-the-shelf instrument. The first and third chapters are mainly concerned with a comprehensive survey of the theory of confocal imaging by the master himself: Tony Wilson. This is supplemented in chapter 5 by a sophisticated treatment of spatial filtering at the confocal pupils. Stephen Hawking was warned, when preparing his popular book *A Brief History of Time* (Bantam, 1988), that every equation halves the readership. If true, this bodes very ill for the popularity of this book, which has 106 equations in Wilson's two chapters alone! Nevertheless, the nontechnical reader who persists with these chapters and takes the derivation of the equations on trust will be rewarded by a thorough knowledge of the principles of confocal microscopy.

Chapter 2, 4, 6 and 7 are concerned mainly with the theory and practice of three-dimensional image representation and optical sectioning in confocal microscopy. Although the first seven chapters describe many of the latest advances and are essential reading for any serious confocal microscopist, it is the second part of the book that, for me, contains a vision of the future. Chapters 8 and 15 discuss the latest developments in phase contrast and interference techniques for the confocal microscope and chapters 9 and 14 discuss the theory and application of Nipkow disc instruments. These confocal microscopes are often considered the poor cousins of stage scanning and mirror scanning instruments but their feature of multiple scanning probes may eventually prove a decisive advantage for high-speed, three-dimensional scanning of living tissue. Taken together, advances in these areas could lead to instruments capable of recording in great detail the full four dimensional space/time structure of the early living embryo using noninvasive optical contrasting techniques. Further speculation on the direction of future developments in cell biology is to be found in chapter 11, whereas the remaining chapters discuss the latest confocal techniques in other major areas of application, including semiconductor metrology and ophthalmology.

In considering the book's potential market, it is easier to say what the book is not rather than what it is. It is certainly not intended as a 'Which Confocal Microscope?' buyer's guide and it is unfortunate for those seeking such advice that Wilson considers it invidious to compare products. Nor is it a practical guide to using confocal microscopes: although much practical advice can be gleaned from it I would suggest (though I hope not invidiously) that biologists will probably find more in the recent *Handbook of Biological Confocal Microscopy* (Plenum, 1990) edited by James Pawley. Aristotle is reputed to have said that 'When a thing has been said once it is hard to say it differently', and it is true that these two books have much in common. Yet I think that

Wilson's book, giving a more thorough and cohesive treatment of the principles of confocal microscopy, will appeal more to designers and developers, whereas users might favour Pawley's book. I am not sure whether either book qualifies as a *vade mecum* but I suspect that the true aficionado of confocal microscopy will not wander far without his or her copy of Wilson. □

Graham Dunn is in the MRC Muscle and Cell Motility Unit, Department of Biophysics, King's College London, 26 - 29 Drury Lane, London WC2B 5RL, UK.

Inscrutable matter

John L. Friedman

Black Holes and the Universe. By Igor Novikov. Cambridge University Press: 1990. Pp.176. Hbk £25, \$39.50; pbk £8.95, \$14.95.

By 1784, it was already clear to the English physicist John Michell that there might be stars large and dense enough that, to escape their surface, one must exceed the speed of light. "Since their light could not arrive at us... we could have no information from sight. Yet, if any other luminous bodies should happen to revolve around them, we might still perhaps from the motions of these revolving bodies infer the existence of the central ones with some degree of probability". Objects from the familiar world of newtonian physics, Michell's black stars are curious but easily understood. More difficult to grasp is their modern incarnation, the black holes of general relativity. Novikov has written, for the scientifically literate layman, an entertaining, and unusually detailed introduction to these realistic black holes and the universe that they apparently pervade.

If gravity were really newtonian, the stellar corpses that Novikov describes would be denser than other stars and dimmer, but neither black nor holes. Light could travel outward from their surface before falling back, and it would be possible to observe them. In fact, with a sufficiently powerful spacecraft, one could visit them and return. But in the real universe, the gravity of a star dense enough to trap its own light has strength to bend inward the path of all its matter. Nucleons, electrons, neutrinos and light fall to a centre, their extent collapsing to the circumference of a speck. Inside the speck, density and curvature are too high for present physics to describe. Outside the

speck space is empty, remembering in its mathematically precise curvature only the mass and rotation of a vanished star. Because the space between speck and horizon is empty, and because any matter crossing the horizon falls inward, 'black hole' is a plausible metaphor. Novikov quotes the poet Marina Katys: "a gulf against the shores of space and time".

Novikov is the latest in a sequence of experts in the field to portray without mathematics the remarkable role that general relativity has come to play in astronomy. The famous co-author (with Y.B. Zel'dovich) of an indispensable text in relativistic astrophysics has given us a more complete picture of black-hole astrophysics than one finds in the similar short books by Robert Wald (*Space, Time, and Gravity: The Theory of the Big Bang and Black Holes*, University of Chicago Press, 1977) and Stephen Hawking (*A Brief History of Time: From the Big Bang to Black Holes*, Bantam, 1988). His discussion of black holes may be less accessible than Hawking's, but Novikov's own pioneering contributions (including the first published prediction with Zel'dovich and Doroshkevich of the 3K cosmic microwave background radiation) more often involve the big bang than black holes. With Marina Katys' poetry, a Soviet song, and his own

John Michell — contemplating the existence of black stars.

anecdotes, Novikov leads us to his truer love. He has a talent for telling stories of scientific discovery, and his portrayal of the young science of physical cosmology is a drama of unexpected brilliance and blindness. □

John L. Friedman is in the Department of Physics, University of Wisconsin, Milwaukee, Wisconsin 5320, USA.

Growing pains

Stuart Sutherland

Human Change Processes: The Scientific Foundations of Psychotherapy. By Michael J. Mahoney. *Basic Books*: 1991. Pp.590. \$45.

SCIENCE has two main aspects, the systematic recording of observations and the construction of rigorous theories to explain them. The subtitle of *Human Change Processes: The Scientific Foundations of Psychotherapy* should constitute an offence under the Trades Descriptions Act: there is nothing in the whole book that remotely resembles either the data or the theories of science. It is the sort of book that brings psychology into disrepute, which is a pity as even on the difficult problem of psychotherapy one or two facts have been established, even if there is still nothing meriting the name of theory.

The book starts with a potted history of philosophy, ranging from Heraclitus (who you will remember never stepped into the same river twice) to Kant. Next, it purports to give a history of 'Scientific Studies of Human Learning and Development'. Although the word 'scientific' keeps popping up in unexpected places, we are treated only to a stream of opinions, many of them banal, such as "David Foulkes studied the dreams of 150 children . . . That research [perish the word] has portrayed dreams as . . . predominantly based in (or drawing upon) what we know and remember". Here is another of Michael Mahoney's insightful conclusions: "To sum up, there is no question that memory and what we experience as mental representations are important phenomena in our lives." He fares no better with the 'theories' he reviews, an activity he describes as 'metatheorising'. "Constructionism has been used with two different meanings: as a portrayal of the organism as an active agent in its own ongoing development, and as a means of highlighting the social contexts that construct and orient our efforts at knowing, communicating, and becoming." When translated into English, this would appear to mean that social scientists believe that although we are influenced by others, we take our own decisions.

Mahoney provides little in the way of argument. He supports his idea that "self examination is important in optimal psychotherapy" by presenting the results of a poll he conducted among psychotherapists: 2,000 psychotherapists cannot be wrong. Or again, he conveys the depressing news that "there are now dozens of distinguishable theories of emotionality, hundreds of volumes devoted to that topic, and tens of thousands of articles dealing with various aspects of human affect." He does not stoop to tell his readers what has been found out about the emotions, perhaps wisely for almost nothing worth knowing has been discovered.



The American bison — about 50,000 remain from an original population of 60 million. Taken from *The Natural History of North America* by Edward Ricciuti (Facts on File, \$35, £18.95).

Having listed the opinions on human development of everyone he can think of and having given them all a prize for their sagacity, if not for their ability to prove their point, he turns in the final section, which has nothing to do with the other two, to the techniques of psychotherapy. Apparently it helps his patients if they look at themselves in a mirror, utter a series of random thoughts, dance, meditate, have massage, play different roles, indulge in 'private dialogue' and write an autobiography (one hopes not for publication). He notes that his patients 'oscillate': remarkably, they are happier at some times than at others. Mahoney never considers such questions as whether psychotherapy is a useless self-indulgence, except possibly for the mentally disordered. The majority of its clientele in America appear to be mentally normal: they just want to 'grow'. Despite Mahoney's assertion to the contrary, it is rare to see anyone over 30 or 40 growing, though they often change for the worse. Nor does he question the desirability of examining the patient's past, which cannot be changed, rather than persuading him to examine his future over which he has some control. Mahoney complains that "Federal funding for research in the social and behavioural sciences has been reduced by 25%". Judging by this book, the American Government has at last got something right.

He has a naive faith in his own success as a therapist. There are now many studies of the outcome of therapy: the sort he espouses, namely dynamic therapy, does no better than a placebo control group in which a sham therapist simply allows the patient to talk and occasionally makes supportive noises. Moreover, despite Mahoney's insistence on the need for training, it has been shown that completely untrained counsellors do precisely as much or as little good as those who have been trained. He is unable to understand that if you have made sacrifices to enter a profession and if you are now making a

decent living out of it, you are likely to think that you are doing some good even if you are not.

Despite his ability to cover nearly 600 pages with verbiage, Mahoney is clearly a very nice man. He actually submits to psychotherapy himself because he feels that only by 'growing' and learning how to 'cope' with his problems can he be of maximum use to his patients. There is fortitude and self-sacrifice indeed.

Occasionally Mahoney says something that needs saying, perhaps after the fashion of the celebrated monkeys who wrote the whole of Shakespeare by scribbling at random. For example, he rightly attacks those psychotherapists who preach to their patients that it is all in their heads on the grounds that "The parameters responsible for the pain of starvation, rape, and oppression are not confined to the nervous system of their victims." Moreover, some of his case histories are interesting. The patient always emerges a better person, but, be warned, the case history is the modern form of fairy story.

Goodness is unfortunately not enough. Even setting aside its lack of argument or evidence, Mahoney's book is extremely pretentious. 'Thinking' becomes 'mentation', 'discussion' 'dialogue', and 'view' 'perspective': all of these words are of course usually prefaced by 'ongoing'. If scholarship were evaluated by the ratio of references to text, Mahoney would be the world's greatest scholar: even the most obvious statements are often supported by up to 40 references.

I hope I have not misjudged this book. There are after all five extremely glowing encomia on the back cover: but then they are all from psychotherapists. I rest my case. □

Stuart Sutherland is in the Department of Experimental Psychology, University of Sussex, Sussex House, Falmer, Brighton. BN1 9QG, UK.

Enzyme catalysis: not different, just better

Jeremy R. Knowles

Where are we in our understanding of enzyme catalysis? The gloomier view is that protein structure and enzyme function are the finely balanced end-products of many weak interactions that can be summed only by massive computing power, and more precise parameterization than we enjoy at present. The cheerier position is that proteins are built on definable principles, and that enzymes use recognizable catalytic devices that will allow us to understand how existing enzymes work and to design new ones. To assess which interpretation is the more realistic, the simple reaction catalysed by triosephosphate isomerase is considered here. This examination illustrates some of the catalytic features of enzymes that are understood, and exposes a few that are not. But overall, the question turns out to have an optimistic answer.

TRIOSEPHOSPHATE isomerase catalyses the chemically simple interconversion of the two triose phosphates, dihydroxyacetone phosphate and *D*-glyceraldehyde 3-phosphate, through the intermediate formation of a *cis*-enediol(ate)¹⁻³ (Fig. 1). The enzyme merely mediates the protonations and deprotonations necessary to effect two enolizations, one from dihydroxyacetone phosphate to the enediol, and the other from glyceraldehyde 3-phosphate to the enediol. When we compare the catalytic power of the isomerase with that of a simple organic base having a $pK_a \sim 7$ (a notional catalyst from the primordial soup), the isomerase is 10^{10} times more effective⁴ (Fig. 2). To improve this primitive catalyst, the kinetic barriers for each of the enolizations must be lowered (Fig. 2, solid arrows), and the enediol(ate) intermediate must be stabilized (Fig. 2, open arrow). How does triosephosphate isomerase do it?

There are, of course, many ways of summarizing the details of an enzyme's active site in terms of devices and motifs, yet in this case a reasonable description of the catalytic apparatus is possible with only three elements: a protein loop that binds and stabilizes the reaction intermediate (fulfilling the open arrow of Fig. 2), and a catalytic base and a catalytic acid to mediate the two enolizations (fulfilling the solid arrows of Fig. 2). What do we know about each of these features?

Intermediate sequestration

Many enzymes have flexible loops, floppy tails, mobile lids, or bivalve-like hinged domains, which appear to close down over the substrate(s) during the chemical changes of the catalytic act⁵⁻¹⁰. Reasons for such sequestration are usually not hard to find: for example, reactive intermediates should not be released into the solvent where they may decompose; stereoelectronics may require a particular substrate conformation to be maintained; water must be excluded; or regiospecificity or stereospecificity demands a tight grip on the reactants. Yet the catalytic value of these conformation changes is often presumed rather than proven.

In triosephosphate isomerase, a loop of ten or eleven amino acids between residues 166 and 176 moves about 7 Å from an 'open' position in the unliganded enzyme into a 'closed' position when there is substrate or inhibitor bound at the active site (Fig. 3a)¹¹⁻¹³. From crystallographic work and molecular dynamics simulations, the conserved residues of this loop (those between residues 168 and 173) seem to form a rigid lid¹⁴ that moves on hinges provided by the unconserved regions on each side. Within the conserved segment, residues 170 to 173 form a looplet that, when the lid is closed, provides a new hydrogen bond to the substrates' phosphate group. As the α carbon of residue 169 is less than 5 Å from that of 174, excision of the four residues of the looplet will not significantly distort the rest of the molecule¹⁵; but this deletion does make the active site lid too small to envelop the bound substrate (Fig. 3b). This mutant protein has

been generated, and turns out to have lost nearly 10^5 -fold in k_{cat} terms. Substrate binding is not seriously impaired, but the enediol phosphate intermediate analogue, phosphoglycolohydroxamate, is bound much more weakly¹⁵. It thus appears that the loop closure preferentially stabilizes the enediol phosphate intermediate, as well as the two transition states that flank it (Fig. 2).

Determination of the full energetics of this loopless isomerase confirms that the mutant enzyme has lost its grip on the enediol phosphate intermediate, this failure being vividly evident from the fact that now only one out of six substrate molecules 'makes it' through to product. In the absence of a proper lid, the enediol phosphate intermediate is lost from the truncated enzyme into solution, where it rapidly decomposes (to give methylglyoxal and P_i). (The ingenuity of the enzyme in binding substrate with the phosphate group in the plane of the enediol, so that this decomposition of the intermediate is stereoelectronically disfavoured in the active site, was first noted by Albert Eschenmoser in 1974). The flexible loop of triosephosphate isomerase thus has two catalytic functions: it ensures an efficient throughput of substrate to product, and it stabilizes the reaction intermediate (solid arrow in Fig. 2), which pulls down the free energy of the two flanking transition states. How the loop manages to be so discriminating, in stabilizing the reaction intermediate while affecting the bound substrate and product to a much lesser extent, will have to await a comparison of the high-resolution structures of the enzyme complexed with substrate and with the intermediate analogue phosphoglycolohydroxamate.

Significantly, the loop that we mutilated contains the amino-acid sequence -A-X-G-X-G-K-X-A- (single-letter notation), which has some similarity to the consensus turn that interacts with phosphate groups in some kinases, many dehydrogenases, *ras* p21, and other nucleotide-binding proteins¹⁶⁻²¹. It seems

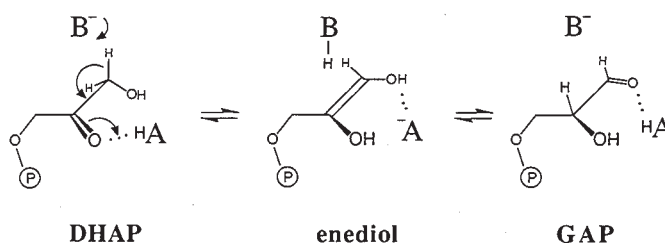


FIG. 1 The reaction catalysed by triosephosphate isomerase. A basic catalytic group B abstracts the pro-*R* proton on C-1 of dihydroxyacetone phosphate (DHAP), assisted by an acidic catalytic group HA, to produce the intermediate enediol. This intermediate then collapses to give the product glyceraldehyde 3-phosphate (GAP) and regenerates the enzyme. The same enediol is produced by the enolization of DHAP or of GAP, and the overall reaction simply puts the two enolizations back-to-back.

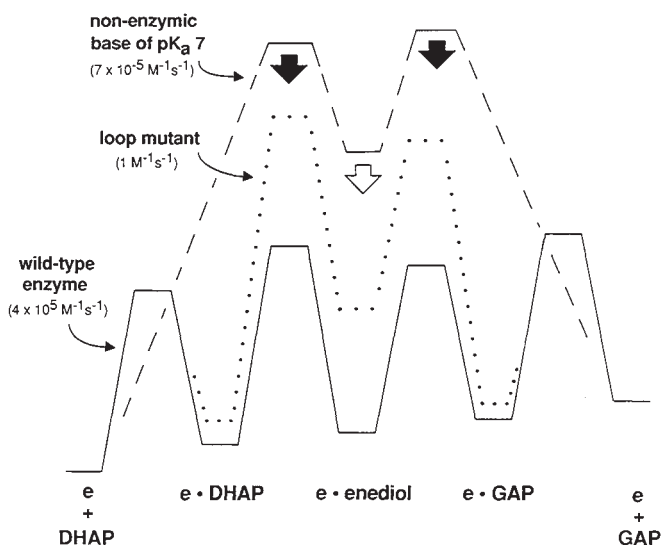


FIG. 2 Free-energy profiles for the isomerization of dihydroxyacetone phosphate (DHAP) and *R*-glyceraldehyde 3-phosphate (GAP), catalysed by wild-type triosephosphate isomerase⁴⁵ (solid line), by the loop mutant enzyme that lacks residues 170–173 (ref. 15) (dotted line), or by a non-enzymic base of $pK_a \sim 7$ (ref. 4) (dashed line). e, Enzyme.

that triosephosphate isomerase has simply recruited, for the catalytic purposes described above, the motif that we have (not altogether frivolously) called the 'phosphate gripper'.

A poised base

From experiments involving kinetics, stereochemistry, chemical modification and X-ray crystallography, we know that the carboxylate of glutamate 165 is the base responsible for the abstraction of the pro-*R* proton of the substrate dihydroxyacetone phosphate (or, for the reverse reaction, the proton at C-2 of *R*-glyceraldehyde 3-phosphate; Fig. 1). From the crystal struc-

tures of both the chicken and the yeast isomerases with either the substrate or a substrate analogue at the active site, it is clear that the oxygens of the carboxylate of Glu 165 are beautifully positioned for the proton abstractions required (Fig. 3c)^{11–13,22,23}. This arrangement is exquisite for four reasons. First, the fit is snug (Fig. 3c), the distance from the nearer carboxylate oxygen to the substrate's C-1 or C-2 being only 2.8 or 3.4 Å. Second, the carboxylate group is bidentate, which means that the proton can be removed either from C-1 of dihydroxyacetone phosphate (from left to right in Fig. 1) or from C-2 of glyceraldehyde 3-phosphate (from right to left in Fig. 1), with minimal motion of the catalytic base. Third, the stereoelectronic requirement for rapid enolization²⁴ (that the abstracted proton lie orthogonal to the enolate plane defined by O-1, C-1, C-2 and O-2; see Fig. 4a) is nicely satisfied by the active site arrangement of triosephosphate isomerase. Fourth, the argument of Gandour²⁵ that a *syn* orbital of a carboxylate ion is much more basic than an *anti* orbital (Fig. 4b), has also been recognized in the configuration shown in Fig. 3c. So, with all the hindsight and all the physical-organic prejudice available, the choice and position of Glu 165 seem ideal. Indeed, when this group is moved by only ~ 1 Å (as in the mutant enzyme where Glu 165 is changed to Asp, for which the crystal structure shows no other significant alterations; E. Lolis and G. Petsko, unpublished results), the catalytic power of the enzyme falls by nearly a thousandfold²⁶. Although it is too early to generalize, it is evident that in this case at least, the positioning of functionality at the active site of the enzyme needs to be quite precise if full catalytic potency is to be realized. Finally, as has been shown for a number of enzymes, complete removal of the catalytic group and its presumed replacement by water (as when Glu 165 is substituted by alanine or glycine) cuts the catalytic activity by more than a million times (K. Liu, unpublished work). The good news for catalyst engineers is that proper placement of appropriate groups in the right environment seems to be enough. The not-so-good news is that this placement must be very precise. (In this regard, it has been argued by several

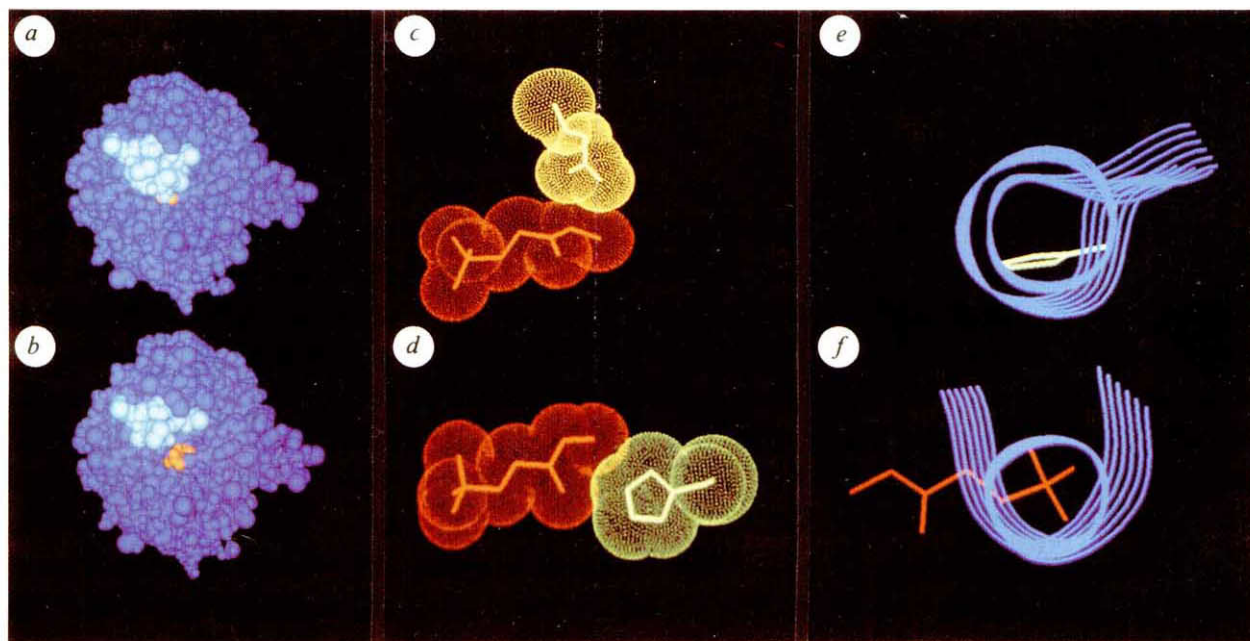


FIG. 3 Catalytic elements of triosephosphate isomerase. *a*, The substrate (orange) is sequestered from solvent by the complete mobile loop (turquoise) (the coordinates are from the crystal structure of the yeast enzyme containing bound phosphoglycolohydroxamate^{22,23}); *b*, the substrate (orange) in the same position as in *a*, cannot be sequestered by a loop (turquoise) from which four amino acids (residues 170–173) have been removed (structure from *a*, after emendation and energy minimization); *c*, shows the snug fit of

substrate (red) and the catalytic base of Glu 165 (yellow) (the dot surfaces are at the van der Waals' radii); *d*, shows the snug fit of substrate (red) and the catalytic electrophile of His 95 (green) (the dot surfaces are at the van der Waals' radii); *e*, showing the α helix (blue) that is aimed at the imidazole ring of His 95 (green); and *f*, shows the α helix (blue) that is aimed at the phosphate group of the substrate analogue phosphoglycolohydroxamate (red).

authors that the catalytic potential of the carboxylate group has rarely, if ever, been realized in model systems, because of the failure to exploit the *syn* orbitals of this group^{27,28}.)

An unexpected electrophile

Early experiments indicated that triosephosphate isomerase also uses electrophilic catalysis to speed the enolization steps, for the carbonyl group of the substrate is considerably polarized on binding to the enzyme^{29,30}. A likely candidate for this electrophilic role is histidine 95, the N ϵ of which lies within 2.9 Å of each of the substrate oxygens on C-1 and C-2 (Fig. 3d). As with the catalytic base (the carboxyl group of Glu 165), the position of the imidazole ring of His 95 with respect to the substrate seems particularly appropriate, and provides for a strong hydrogen bond to the carbonyl oxygen of either of the substrates dihydroxyacetone phosphate or glyceraldehyde phosphate. There is even the seductive possibility (originally formulated by Jules Rebek) that the proton transfers are mediated merely by rotation of the appropriate side chain: the bidentate carboxyl group for the manipulation of carbon-bound protons, and the bidentate imidazole group for handling the oxygen-bound protons. Mutagenesis experiments have confirmed the importance of His 95 (conversion of this residue to glutamine gives an enzyme that is 100 times less active; conversion to asparagine cuts the activity by 10⁴; refs 31–32), and the electrophilic and proton transfer roles of the imidazole ring have been further illuminated by infrared and X-ray crystallographic work on these mutant enzymes³³. Indeed, when glutamine replaces His 95, the enzyme—deprived of its catalytic electrophile—finds a subtly different mechanistic pathway to achieve the interconversion of the two substrates at 1% of the wild-type rate³¹. This mutant apparently uses Glu 165 opportunistically to effect all the proton transfers, namely those to and from both carbon and oxygen centres.

On this basis, the reader would be forgiven for concluding that Fig. 3c and d adequately describes a precisely positioned

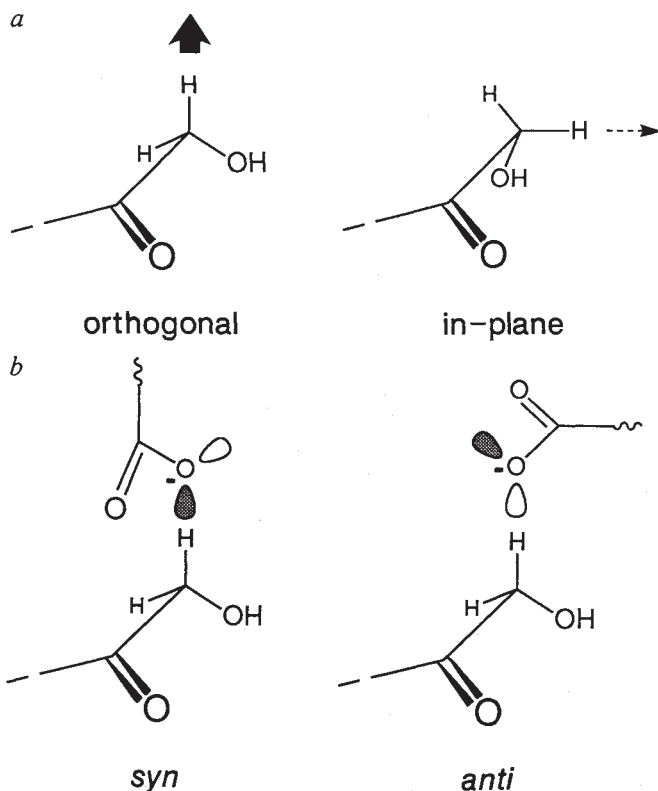


FIG. 4 Optimum geometry for enolization by a carboxylate base. *a*, The proton abstracted should lie orthogonal to the plane of the enolate²⁴; and *b*, a carboxylate *syn* orbital (shaded) is believed to be more basic than an *anti* orbital (open)²⁵.

general base (the carboxylate of Glu 165) and a well-placed general acid (the imidazolium of His 95), which are poised for 'push-pull' catalysis. Yet this is not how things are, and our understanding is not complete. The crystal structure of the isomerase indicates from the pattern of hydrogen bonds at the active site^{11–13} that the imidazole ring of His 95 is not protonated at neutral pH; this implies that the enzyme uses imidazole rather than the much more electrophilic imidazolium as its acid component. We have investigated this using ¹⁵N NMR, which is particularly suitable for such purposes³⁴. First, all histidines in the yeast enzyme except that at position 95 were replaced with glutamine, and the resulting (fully active) enzyme was isolated, selectively enriched with ¹⁵N-histidine. Using ¹⁵N NMR, we found that the pK_a of His 95 is lowered from 6.7 in the denatured protein to less than 4.5 in the native enzyme, and we observed the formation of a new hydrogen bond (presumably to substrate) from the neutral imidazole ring (P. Lodi, unpublished), indicating that the enzyme does use the less acidic neutral imidazole as its electrophile. Why, when everything else about this enzyme seems chemically so reasonable, it should fail to use the more powerful electrophile, remains a puzzle. Yet, so as not to end in a minor key, let us look at one other feature of the active site.

Well-aimed helices

Hol has published a splendid review³⁵ in which he develops the idea that the formal dipole associated with a classical α helix modulates the properties of groups at the helix termini. Whether the observed effects are due to the electric field of a helix dipole or merely to the local fields of the unsatisfied carbonyl groups and —NH— groups at each end of the helix³⁶, is less important than the effectiveness of the device, which is well established^{37,38}. Triosephosphate isomerase has two conserved α helices that are directed towards the active site. One helix is trained on His 95, the imidazole ring of which is at the positive end (Fig. 3e). The aim is precise and, even if we cannot understand why the enzyme should want to lower the pK_a of His 95 to less than 4.5 and so ensure a neutral imidazole, the arrangement of this stretch of helix accounts for the perturbation very well.

The second helix that points into the active site has its positive

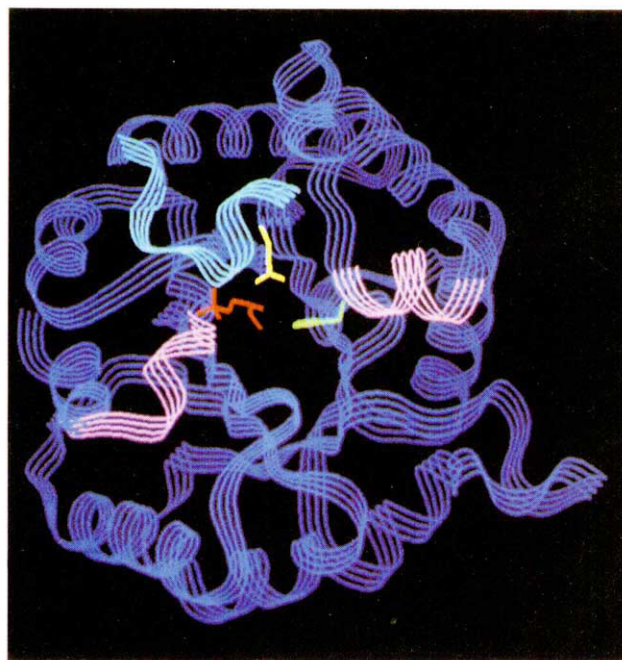


FIG. 5 One subunit of triosephosphate isomerase, with bound substrate (red), Glu 165 (yellow), His 95 (green), the two α helices that are aimed at the active site (pink), and the mobile loop that closes over the bound substrate (light blue).

end aimed at the phosphate group of the substrates (Fig. 3f). This arrangement too, is attractive, because there is only one positively charged group at the active site (Lys 12), and yet the substrates bind as their phosphate di-anions³⁹. The arrangement shown in Fig. 3f provides two hydrogen bonds from main chain —NH— groups (of Gly 232 and Gly 233) to complete the neutralization of charge on the phosphate group.

If a single message emerges after all of the mechanistic and structural scrutiny of triosephosphate isomerase, it is one of precision⁴⁰, most evident from Fig. 3c–f. That the fit of substrate (or, more strictly, transition state^{41–43}) to active site is exact has been clear for nearly a century⁴⁴, but there are other elements

of precision that perhaps we have not fully appreciated. Indeed, one may speculate that the relatively large size of enzymes (compared with most of their substrates) may derive from the need for a matrix that positions their functional groups, focuses their helices and anchors the ends of their mobile loops (Fig. 5). And if our bioorganic effort and chemical mimicry continue to fall short of the catalytic potency of today's enzymes, at least we know that nature has been adjusting and refining for rather longer than we have⁴⁵. □

Jeremy R. Knowles is in the Departments of Chemistry and Biochemistry, Harvard University, Cambridge, Massachusetts 02138, USA.

1. Rieder, S. V. & Rose, I. A. *J. biol. Chem.* **234**, 1007–1010 (1959).
2. Rose, I. A. *Brookhaven Symp. Biol.* **15**, 293–309 (1962).
3. Maister, S. G., Pett, C. P., Albery, W. J. & Knowles, J. R. *Biochemistry* **15**, 5607–5612 (1976).
4. Richard, J. P. *J. Am. chem. Soc.* **106**, 4926–4936 (1984).
5. Adams, M. J. et al. *Proc. natn. Acad. Sci. U.S.A.* **70**, 1968–1972 (1973).
6. Clarke, A. R. et al. *Nature* **324**, 699–702 (1986).
7. Ludwig, M. L. et al. *Proc. natn. Acad. Sci. U.S.A.* **57**, 511–514 (1967).
8. Bennett, W. S. Jr & Steitz, T. A. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4848–4852 (1978).
9. Pickover, C. A., McKay, D. B., Engelman, D. M. & Steitz, T. A. *J. biol. Chem.* **254**, 11323–11329 (1979).
10. Winn, S. I., Watson, H. C., Harkins, R. N. & Fothergill, L. A. *Phil. Trans. R. Soc.* **B293**, 121–130 (1981).
11. Banner, D. W. et al. *Nature* **255**, 609–614 (1976).
12. Alber, T. C. et al. *Phil. Trans. R. Soc.* **B293**, 159–171 (1981).
13. Lolis, E. & Petsko, G. A. *Biochemistry* **29**, 6619–6625 (1990).
14. Joseph, D., Petsko, G. A. & Karplus, M. *Science* **249**, 1425–1428 (1990).
15. Pompliano, D. L., Peyman, A. & Knowles, J. R. *Biochemistry* **29**, 3186–3194 (1990).
16. Müller, W. & Amons, R. *FEBS Lett.* **186**, 1–7 (1985).
17. Wierenga, R. K., Terpstra, P. & Hol, G. J. *J. molec. Biol.* **187**, 101–107 (1986).
18. Seeburg, P. H., Colby, W. W., Capon, D. H., Goeddel, D. V. & Levinson, A. D. *Nature* **312**, 71–75 (1984).
19. Reinstein, J., Brune, M. & Wittinghofer, A. *Biochemistry* **27**, 4712–4720 (1988).
20. Redfield, A. G. & Papastavros, M. Z. *Biochemistry* **29**, 3509–3514 (1990).
21. Saraste, M., Sibbald, P. R. & Wittinghofer, A. *Trends biol. Sci.* **15**, 430–434 (1990).
22. Davenport, R. C. thesis, MIT, Cambridge, Massachusetts (1986).
23. Alber, T. C. et al. *Cold Spring Harbor Symp. quant. Biol.* **52**, 603–613 (1987).
24. Corey, E. J. & Sneen, R. A. *J. Am. chem. Soc.* **78**, 6269–6278 (1956).
25. Gaudour, R. D. *Bioorg. Chem.* **10**, 169–176 (1987).
26. Raines, R. T., Sutton, E. L., Straus, D. R., Gilbert, W. & Knowles, J. R. *Biochemistry* **25**, 7142–7154 (1986).
27. Zimmerman, S. C. & Cramer, K. D. *J. Am. chem. Soc.* **110**, 5906–5908 (1988).
28. Huff, J. B., Askew, B., Duff, R. J. & Rebek, J. J. *Am. chem. Soc.* **110**, 5908–5909 (1988).
29. Webb, M. R. & Knowles, J. R. *Biochem. J.* **141**, 589–592 (1974).
30. Belasco, J. G. & Knowles, J. R. *Biochemistry* **19**, 472–477 (1980).
31. Nickbarg, E. B., Davenport, R. C., Petsko, G. A. & Knowles, J. R. *Biochemistry* **27**, 5948–5960 (1988).
32. Blacklow, S. C. & Knowles, J. R. *Biochemistry* **29**, 4099–4108 (1990).
33. Komives, E. A. et al. *Biochemistry* (in the press).
34. Bachovchin, W. W. *Biochemistry* **25**, 7751–7759 (1986).
35. Hol, W. G. J. *Prog. biophys. molec. Biol.* **45**, 149–195 (1985).
36. Tidor, B. & Karplus, M. *Biochemistry* (in the press).
37. Perutz, M. F., Gronenborn, A. M., Clore, G. M., Fogg, J. H. & Shi, D. T.-b. *J. molec. Biol.* **183**, 491–498 (1985).
38. Sali, D., Bycroft, M. & Fersht, A. R. *Nature* **335**, 740–743 (1988).
39. Belasco, J. G., Herlihy, J. M. & Knowles, J. R. *Biochemistry* **17**, 2971–2978 (1978).
40. Blow, D. M. & Smith, J. M. *Phil. Trans. R. Soc.* **B272**, 87–97 (1975).
41. Haldane, J. B. S. in *Enzymes* 182 (Longmans, London, 1930).
42. Pauling, L. *Chem. Eng. News* **24**, 1375–1377 (1946).
43. Leatherbarrow, R. J., Fersht, A. R. & Winter, G. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7840–7844 (1985).
44. Fischer, E. *Ber. dtsh chem. Ges.* **27**, 2985–2993 (1894).
45. Albery, W. J. & Knowles, J. R. *Biochemistry* **15**, 5627–5631 (1976).

ACKNOWLEDGEMENTS. I thank, among many collaborators on the isomerase problem, R. T. Raines, A. Peyman, D. L. Pompliano, E. A. Komives, K. Liu, P. Lodi, E. Lolis and G. Petsko, upon whose contributions much of this account is based. Computer graphics were by P. Lodi and D. Pompliano. Supported by the General Medical Sciences Institute of the NIH and by the NSF.

ARTICLES

Geodetic determination of tectonic deformation in central Greece from 1900 to 1988

H. Billiris*, D. Paradissis*, G. Veis*, P. England†, W. Featherstone‡, B. Parsons†, P. Cross‡, P. Rands‡, M. Rayson‡, P. Sellers‡, V. Ashkenazi§, M. Davison§, J. Jackson|| & N. Ambraseys¶

* Department of Surveying, National Technical University, Athens 15773, Greece

† Department of Earth Sciences, University of Oxford, Parks Road, Oxford OX1 3PR, UK

‡ Department of Surveying, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, UK

§ Institute of Engineering Surveying and Space Geodesy, University of Nottingham, University Park, Nottingham NG7 2RD, UK

|| Department of Earth Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EQ, UK

¶ Department of Civil Engineering, Imperial College, Imperial College Road, London SW7 2BU, UK

The Global Positioning System has been used to measure the relative displacements of fifteen monuments in a hundred-year-old triangulation network spanning part of the Aegean extensional basin. These displacements reflect the tectonic deformation of the region over the past century, showing more than one metre of north-south extension across the network. The crust in this region appears to contain a few slowly deforming blocks separated by more rapidly deforming zones.

At the same time that plate tectonics was being used successfully to describe a wide variety of geological observations, it was realized that the axioms of plate tectonics do not apply to parts of the continents (see, for example, ref. 1). Deformation is distributed over wide areas on the continents, rather than being concentrated on narrow boundaries as in the oceans. A key aim of tectonics is to determine the most appropriate and economical description of this distributed deformation; without such a description, it is difficult to make progress on studies of the dynamics of the deformation (for example, ref. 2). Most of the quantitative data on active continental deformation come from earthquake seismology and satellite imagery^{3–6}, but this approach is limited. Earthquake focal mechanisms cannot pro-

vide information about rotations, so only the symmetrical part of the deformation field (the strain) can be determined from the seismology; the magnitude of the strain estimated from earthquakes depends on their scalar moments and, for earthquakes in the early part of this century, these are uncertain by about a factor of two; finally, it is known that, in some areas, a large fraction of the deformation occurs without giving rise to earthquakes⁵.

For these reasons, it is desirable to form an independent estimate of the rates of deformation in regions of distributed continental deformation. Such estimates can come from geodesy. If benchmarks are implanted in bedrock and their positions are determined precisely at suitably spaced intervals of time, changes in their relative positions may be measured, and hence crustal deformations can be inferred. The recent development of Global Positioning System (GPS) of navigation yields, as a by-product, a tool well suited for such precise positioning measurements.

GPS is based on the simultaneous determination of the distances between a ground receiver and a number of (usually four or more) satellites. In most applications these distances are determined by measuring the time of travel of packets of binary code that modulate two carrier waves that are emitted by the satellites with frequencies of 1,575.42 and 1,227.60 MHz (wavelengths of about 20 and 25 cm respectively). These packets of code cover many cycles of the carrier waves and carry data that allow the immediate determination of position with accuracies of around 100 m and 10 m (for civilian and military users respectively). Geodesists have found, however, that by measuring the phase of the signal (with an accuracy of about 1°, equivalent to about a millimetre), and by using what are essentially interferometric techniques, much higher relative positioning accuracies can be achieved. The technique suffers from several sources of uncertainty (see below), which make millimetric accuracy a rare achievement, but in favourable circumstances GPS is now capable of determining relative positions with centimetric accuracy over distances as great as 1,000 km (see, for example, refs 7, 8).

GPS is particularly attractive because it gives results in a stable global coordinate system (currently the World Geodetic System 1984—WGS84) and so all elements of the displacement gradient tensor are potentially estimable, including the rotation elements that are inaccessible to seismology.

In practice, deformation of the Earth's surface is so slow that several years must elapse before crustal movements (the signal)

can reliably be distinguished from measurement uncertainties (the noise). Fortunately, in the course of map making, surveyors have covered many parts of the globe, including some that are tectonically active, with classical triangulation networks. In a few cases the original measurements are published and are of high quality, the original benchmarks still exist, and the network is sufficiently old that there has been a resolvable strain since it was installed. In these cases, it is possible to obtain an immediate determination of crustal motions that took place over decades or more.

Between 1890 and 1900 a first-order triangulation network was established across an area that covers what is now Southern and Central Greece (Fig. 1)⁹. Several of the monuments from the 1890s network have survived, protected in part by their remote locations. In September 1988 we re-occupied fifteen of the surviving sites using GPS receivers (Fig. 2) and were able to determine the relative displacements of those sites over the intervening ninety years. Because the seismic activity of the region over the same time interval has been thoroughly studied^{5,10,11}, it is possible to compare directly the geodetic and seismological estimates of relative displacement.

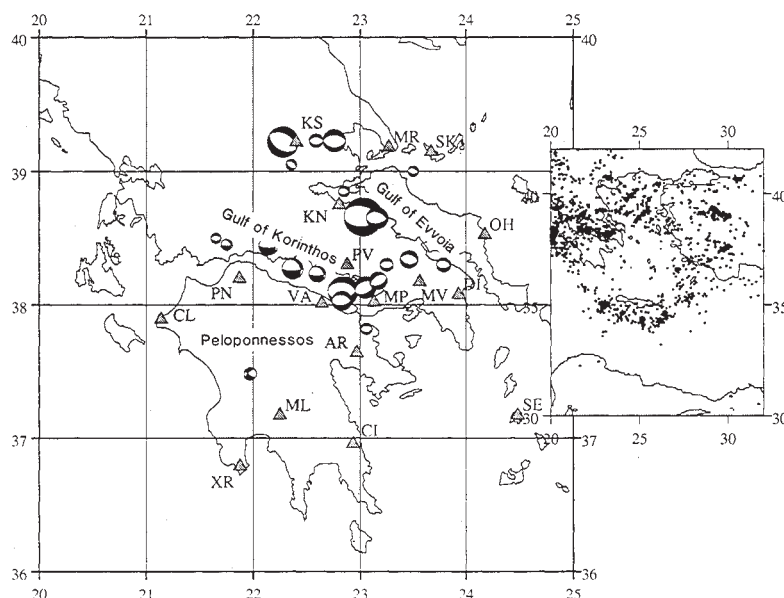
Coordinates of the 1988 GPS network

In addition to the fifteen stations of the 1890s network, two satellite-laser-ranging (SLR) sites of the Wegener/MEDLAS project, at Chrissokellaria (XR) and Dionysos (DI) (Fig. 1), were also included in the GPS network. Observations were made with 9 dual-frequency TI-4100 receivers over an 8-day period, of which seven were observation days, one instrument remaining at Dionysos throughout the survey. Each station was observed for at least two days, and successive days of observation had at least four stations in common.

GPS phase observations, and their processing, suffer from a number of sources of error, of which the most significant are probably those resulting from refraction of the signals in the ionosphere and troposphere, from imprecise knowledge of the positions of the satellites, and from the receivers' losing track of the phase of the incoming signals because of physical obstruction or electrical interference.

We made measurements of phase on the two GPS carrier frequencies and corrected for the influence of ionospheric refraction by the conventional means, which take advantage of the fact that the phase delay in the ionosphere is inversely proportional to the frequency of the signal. We could not correct for refraction in the troposphere because we had no reliable data

FIG. 1 Locations of the sites in the geodetic network discussed here, and fault-plane solutions for all earthquakes with magnitude $M_s \geq 5.8$ that occurred between 1890 and 1988 within that network¹¹. Fault-plane solutions were obtained from teleseismic waveform observations where possible; failing this, the strike, dip and rake were estimated from surface faulting associated with the earthquake. In the absence of these direct observations, the fault-plane solutions of nearby events, or the orientation of nearby active faulting, were used. The areas of the circles containing the fault-plane solution are proportional to the seismic moments of the earthquakes. Triangles denote sites common to the 1890s and 1988 networks as well as the two SLR sites of the 1988 network. The sites are labelled with abbreviations: KS, Kassidiaris; MR, Makra Rachi; SK, Skopelos; KN, Knimis; OH, Ochtonia; PV, Palaiovuna; MV, Megalo Vuno; DI, Dionysos SLR station; PN, Panachaikon; VA, Valtza; MP, Makriplaji; CL, Chlemutsi; AR, Arachneon; XR, Chrissokellaria SLR station; ML, Malevon; CI, Chionovuni; SE, Serifos. Inset shows the regional setting; small dots give the locations of earthquakes with $m_b > 5$ for the map area taken from the USGS Preliminary Determination of Epicenters.



Central Greece 1893 - 1981

on the amount of water in the atmosphere, such as could be provided by weather balloons or water-vapour radiometers. Consequently, uncertainty at the one- or two-centimetre level remains from this source. In our analysis, we have used positions of the GPS satellites as given by their own broadcast estimates of their orbits. Orbit errors of about 20 metres lead to scale uncertainties of around 1 part in 10^6 . These errors may be reduced either by employing the more precise orbits released later by the US Department of Defense, or by constraining the orbits with GPS phase data collected (at the same time as the survey) at stations of known position. The 1988 coordinates of our network will be re-determined later using precise orbits, but the coordinates reported here are based on the broadcast orbits and so have uncertainties of one or two centimetres from this source.

Because of the complexities in treating the data, it has become standard geodetic practice to process important GPS campaigns with more than one software package. Our data were processed by two independent methods.

The first method used the Bernese V3.1 package¹². The data were processed in seven sets, corresponding to the seven days of observation. The solution for each day yielded, for each station observed that day, three-dimensional cartesian coordinates in the WGS84 system. These coordinates were derived as differences from the position of the Dionysos SLR station. An estimate was obtained for the covariance of the coordinates. The seven days of data (coordinate differences and their covariance matrices) were adjusted¹³ to provide a single estimate for the coordinates of all points, based on the fixed coordinates of Dionysos. A data validation was then carried out that compared the coordinate differences obtained from the three-dimensional adjustment with those generated by the separate daily solutions. The process involved testing based on the tau (τ) distribution¹⁴ and produced the conclusion that there were no significant inconsistencies in the data. The standard deviations of the results, expressed in terms of their horizontal components, are 0.03 m in latitude and 0.04 m in longitude, with respective maximum values of 0.06 m and 0.07 m.

An analysis was also carried out using the MAGNET-101 V7.21 software¹⁵. Although this program is not designed for very-high-precision applications, it was considered adequate to afford an independent confirmation that the results do not contain errors that would invalidate our analysis of the horizontal strain. Again, the analysis involved two stages. In the first, individual baselines and their covariances were determined from the MAGNET software, and in the second stage they were adjusted to give a single estimate of the coordinates¹³. Initial individual baseline estimates from the MAGNET program were then compared with the adjusted baseline values. Where an individual estimate differed significantly from the adjusted values for that baseline, and there were sufficient additional observations of the same baseline to identify the discrepant estimate as an outlier, it was deleted from the final adjustment. In the final adjustment of the baselines determined using the MAGNET software, there were 135 measurements of a network of 44 baselines, of which 15 were measured only once. The standard deviations of the final results are between 0.02 and 0.04 m. Note that this computation is not rigorous in that (unlike the computation based on the Bernese software) it does not take account of correlation between simultaneously observed baselines.

The Bernese and MAGNET solutions have been compared by computing relative motion vectors of the stations (1890s to 1988) for each solution. The differences in the lengths of the vectors determined by these two methods is small: the maximum difference in length is 0.06 m and the average is around 0.02 m, with the MAGNET solution generally being the longer. The maximum difference in orientation of the relative motion vectors is only 10° , with an average around 3° . These differences are

consistent with the uncertainties in position for the Bernese solutions quoted above.

The uncertainties in the Bernese solutions are equivalent to an average horizontal precision of around 0.4 p.p.m. between any pair of stations in the network. Because these coordinates are obtained by the more rigorous procedure, they are the ones referred to as the 1988 coordinates in the discussion below. It must be emphasised, however, that either set of GPS solutions would lead to the same conclusions about the deformation. The GPS results are an order of magnitude more precise than those determined from readjustment of the 1890s observations (below). In turn, the uncertainties in the 1890s coordinates are small compared with the total displacement across the network (Fig. 2).

Coordinates of the 1890–1900 network

The basis for the geodetic determination of the strain of the region is two sets of coordinates for fifteen monuments of the 1890s survey that are separated in time by ninety years. The coordinates of the monuments in September 1888 are expressed in the WGS84 coordinate system. The published observations from the 1890s survey consist entirely of theodolite directions (502 angles involving 93 stations⁹). These are horizontal angles from an arbitrary origin; there is, therefore, no information on the position, orientation or scale of the original network, and assumptions must be made about these before we can compare modern determinations of position with the original observations.

To choose the scale of the original network we make the assumption that the length of a particular baseline in the Peloponnessos (Panachaikon–Chionovuni in Fig. 2a) has not changed between the two surveys. This assumption has some independent support from the absence of large earthquakes in the central Peloponnessos (Fig. 1) and seems to be justified *a posteriori* by the small displacements determined for other sites in the Peloponnessos (Fig. 2a; also see below). The remaining assumptions, to give orientation and translation of the original network with respect to the modern one, are completely arbitrary. Here we discuss mainly those coordinate differences obtained by assuming that the same pair of stations used to fix the scale (PN and CI) did not move with respect to their modern positions (Fig. 2a); but a second, equivalent, representation is shown in Fig. 2b.

Positional uncertainties of up to 50 cm were found for the readjustment of the 1890s network, but the typical uncertainty in the area reoccupied by GPS in 1988 was ~ 10 –40 cm. The mean observational standard error for directions was 0.55 seconds of arc.

Horizontal relative displacements 1890–1988

Horizontal relative displacements are calculated from the difference between the coordinates of the 1988 GPS network and those derived from the adjustment of the 1890s network, as described above. Figure 2a shows horizontal relative displacements calculated assuming Panachaikon and Chionovuni to be fixed to their 1988 positions. Note that these displacements contain no information on translation or rotation of the network as a whole. To emphasize this point, Fig. 2b shows displacements calculated under different assumptions, which are described in the caption.

The assumption made in deriving the scale of the network—that the length of the line joining Panachaikon and Chionovuni did not change between the two surveys—seems to be justified, because the calculated displacements of three other points in the Peloponnessos (Valtsa, Arachnaeon and Malevon) are less than the uncertainties in their original positions (about ten centimetres). There has, therefore, been no resolvable shear strain of the central Peloponnessos over the past 90 years. (The other point on the Peloponnessos (Chlemutsi) seems to have been displaced by about 70 cm (Fig. 2). The triangulation point

is, however, on top of the wall of an old castle; the displacement may be due either to tectonic activity or to settling of the fabric of the castle.) Formally, there could still be an error in the scale of the network, if there had been a purely isotropic dilatation or contraction of the region in the time interval between measurements, but such a strain seems unlikely on geological grounds.

Despite the indeterminacy in translation and rotation of the network as a whole, the quantities of geological interest—the relative displacement gradients (subject to an unknown rigid-body rotation)—can be determined from coordinate differences within the network between the 1890s and the present.

Distribution of deformation

Uncertainties in the original (1890) coordinates are between 10 and 40 cm, and the baselines are 50–150 kilometres long, so strains smaller than about 2×10^{-6} cannot be detected. Figure 3 shows the differences between the horizontal displacements for each pair of points in the network. The shading distinguishes those displacements that are greater than the uncertainties in the original positions from those that are not resolvable, given the uncertainties. There are three regions in which the sites have not moved resolvably with respect to each other. The first group contains five sites within the Peloponnessos: Panachaikon (PN), Valtza (VA), Arachneon (AR), Chionovuni (CI) and Malevon (ML). The very small (~ 10 cm) relative displacements between these points are consistent with the absence of major seismic activity in the central Peloponnessos over the last hundred years (Fig. 1). There are, nevertheless, several active normal faults within the Peloponnessos (see, for example, ref. 16). Most of those faults lie close to the edge of or, like that which slipped in the Kalamata earthquake of September 1986 (22° E, 37° N), outside our network. Although there is a report¹⁸ of Sparta being

destroyed by an earthquake in the fifth century BC, this event need not have been large, and may also have had its epicentre outside the region covered by our network. It appears from the measurements reported here that the Peloponnessos is being strained at least ten times more slowly than the more seismically active region to the north of the Gulf of Korinthos.

The second group contains four sites immediately north of the Gulf of Korinthos: Palaiovuna (PV), Megalo Vuno (MV), Makriplaji (MP) and Ochtonia (OH). Relative displacements between these points are discernably larger than between those of the previous group (Fig. 3). Unlike the Peloponnessos, this region has experienced several large earthquakes since 1890 (Fig. 1); these could account for displacements of a few tens of centimetres between the stations (see Fig. 1 and Tables 1 and 2 of ref. 11).

The third group lies in the northernmost part of the network. Kassidiaris (KS), Makra Rachi (MR) and Knimis (KN) have not moved resolvably with respect to each other and since 1890 there have been no large earthquakes within the triangle described by these points, although there have been several immediately to the north of this part of our network (Fig. 1).

The displacement of the monument on Serifos with respect to both the Peloponnessos and central Greece is unexpected, given the low level of seismic activity in this part of the Aegean (Fig. 1). Until other 1890s points near Serifos are re-occupied, however, we cannot assess the significance of this displacement.

Comparison with seismological estimates

Approximately 1.1 m of extension has occurred across the network in the 90–100-yr interval between the surveys, as indicated by the displacements of the northernmost sites in Table 1. The portion of the network in which the extension occurs (to

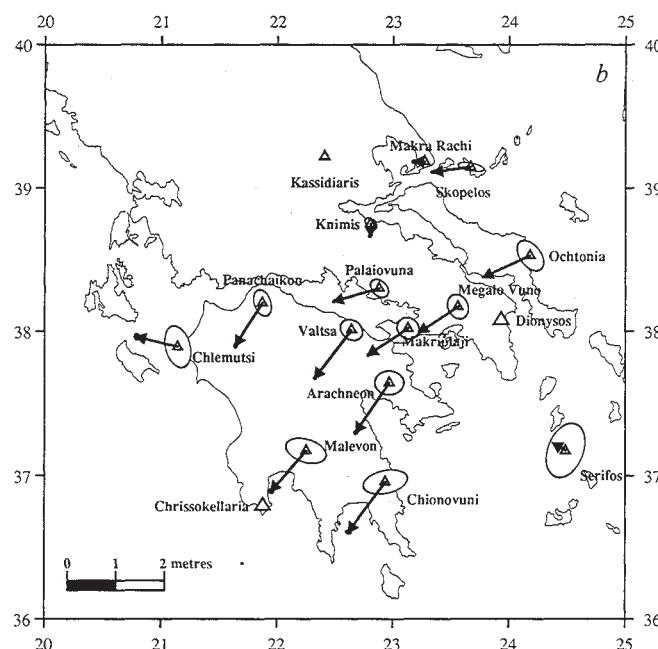
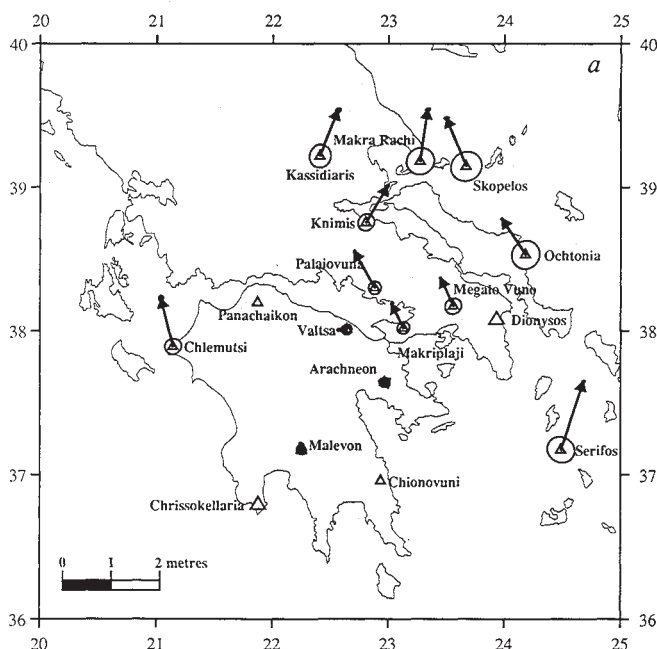


FIG. 2 *a*, Horizontal relative displacements calculated from coordinates derived for the 1890s triangulation network and for the 1988 GPS network, assuming Panachaikon and Chionovuni to be fixed in their 1988 positions. Ellipses centred around the sites (triangles) give the uncertainties associated with the readjustment of the 1980s network. Ellipses corresponding to the positions of the 1988 network are plotted at the head of the displacement vectors but, at the scale used for the displacements, are too small to be clearly visible. Larger triangles, without arrows or ellipses, are (modern) SLR sites of the Wegener/MEDLAS project. The bar at the lower left shows the scale for the arrows and ellipses. *b*, Horizontal relative displacements

calculated with orientation, translation and scale given by the following assumptions: the scale is fixed by holding the length of the line between Panachaikon and Chionovuni to its 1988 length; Kassidiaris is held to its 1988 position; and the line joining Kassidiaris to Makra Rachi is held at its 1988 orientation. These conditions reflect the presumption that the most northerly part of the network has been displaced least with respect to Eurasia in the last 100 years. Apart from small differences in coordinates of the 1890s points owing to the different ways that the uncertainties are distributed during the readjustment, these displacements are equivalent to those in *a* with the addition of a translation and a rotation.

the north of the Peloponnessos) is about 120 kilometres across. The extensional strain is therefore $\sim 10^{-5}$, which corresponds to a strain rate of 10^{-7} yr^{-1} , or $3 \times 10^{-15} \text{ s}^{-1}$.

The displacement accounted for by earthquakes over the same time interval has been calculated by Ambraseys and Jackson¹¹, using the method of Kostrov¹⁶. The major uncertainties in the calculation lie in the estimates of the scalar moments, M_0 , of the earthquakes, which are directly related to the seismic portion of the strain¹⁶. For most of the earthquakes, M_0 was obtained by using an empirical relationship¹¹ between M_s , the magnitude of the earthquake, and M_0 . Uncertainties of ± 0.25 in M_s were used to derive maximum and minimum estimates of the scalar moments. An additional uncertainty arises because the relationship between M_s and M_0 appears to exhibit regional variations¹⁷. The estimates of displacement from seismicity in Table 1 were obtained using the regional relationship that is suggested (ref. 17; see also ref. 11, Fig. 5) to be more appropriate for the Aegean. For further discussion of these uncertainties see references 1, 3, 4 and 13.

The probable¹¹ value of the displacement across this region that is accounted for by earthquakes with $M_s > 5.8$ is 45 cm (Table 1). The slip in all earthquakes with $M_s < 5.7$ increases this quantity by about 50% (ref. 11) to about 70 cm. Extreme values of displacement, allowing for systematic biases of ± 0.25 in M_s , might be as low as 35 cm or as high as 125 cm. Thus the seismic strain between the 1890s and 1988 appears to be about 60% of the geodetic strain, although it may be as high as the geodetic strain if the seismic moments of the earthquakes have been systematically underestimated¹¹.

Comparison with other observations

Arguments based on the regional plate tectonic setting of Central Greece suggested to McKenzie^{10,19} that the Peloponnessos is moving southwestward with respect to northern Greece. Yet the slip vectors in the earthquakes in central Greece are almost entirely north-south⁵. McKenzie and Jackson²⁰ suggested that these observations could be accommodated if the deformation in Central Greece—in the region of this network—were a combination of normal faulting and clockwise rotation of crustal blocks that are separated by these faults. The available palaeomagnetic data^{21,22} suggest that there has indeed been such clockwise rotation over the past few million years.

We would like to be able to compare the observations of slip vectors and rotations with the geodetic determinations of displacements. Unfortunately there is no information about the

TABLE 1 Relative displacements across the network between 1890 and 1988

Site	Geodetic displacement (cm)		
	E-W	N-S	Magnitude
Kassidiaris	39.0	95.1	102.8
Makra Rachi	18.4	108.1	109.7
Skopelos	-32.1	103.5	108.4

	Seismic displacement (cm)		
	E-W	N-S	Magnitude
Minimum	0.7	25.4	25.4
Probable	1.2	45.1	45.1
Maximum	2.2	82.5	82.5

Geodetic estimates are from the relative-motion vectors displayed in Fig. 2a. Seismic displacements are from ref. 11 (see text). In comparing the two estimates of displacement, it must be remembered that the 1890s triangulation contained no information on orientation, and therefore that only the magnitudes of the displacements are directly comparable.

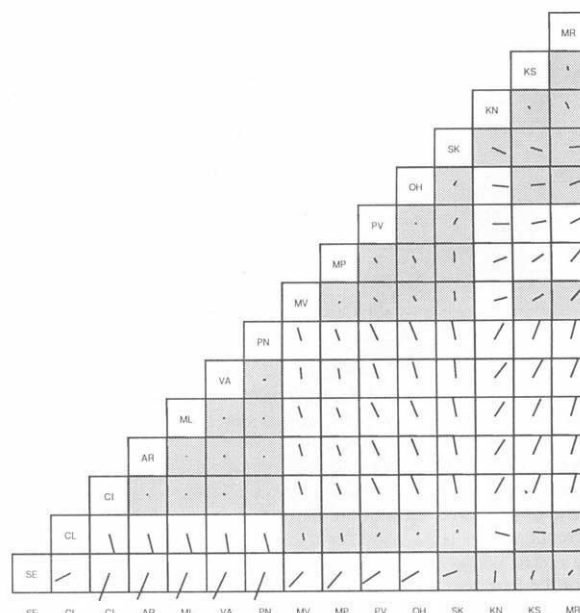
original orientation of the network and, unlike the estimation of the scale, there is no testable assumption that we can make to estimate that original orientation. A simple assumption is that the most northerly part of the network has been rotated least with respect to stable Eurasia. Figure 2b shows the displacements calculated under this assumption.

The general southwesterly movement (Fig. 2b) of the Peloponnessos with respect to northern Greece is consistent with McKenzie's^{10,19} suggestion. Furthermore, the northwesterly movement of the points immediately north of the Gulf of Corinth with respect to the Peloponnessos (Fig. 2a, b) is consistent with a combination of north-south extension and clockwise rotation, as suggested by McKenzie and Jackson²⁰. There are, however, too few points in this survey, and they are placed too far apart, for us to determine whether there is pervasive block rotation of the kind envisaged in ref. 20. It seems clear, however, that if the central Peloponnessos is accepted as a single rigid block, it cannot be rotating appreciably with respect to northern Greece.

Conclusion

It is inevitable that, in a geodetic study based on a network almost a century old, there should be uncertainty about the

FIG. 3 Each square in the diagram shows the displacement of the station given at the head of the column relative to the station at the left of the row; the vector depicting the displacement has its origin at the centre of the square. Squares are white when the magnitude of the relative displacement is more than 2 standard deviations of the error in the relative displacement, and grey when it is less than this. Groups of grey squares over several neighbouring sites provide an indication of areas that may have moved as rigid blocks over the past 100 years. Abbreviations the same as in Fig. 1.



quality of the original data and monumentation. The adjustment, described above, of the published observations of the 1890s survey demonstrates their high quality. (A standard deviation of 0.55 seconds of arc corresponds to one part in 10^6 , which is close to the limit of precision in classical surveys.) The apparent relative horizontal displacements within the central Peloponnese (Fig. 2a) are of the order of 10 cm over baselines 100–150 km long: one part in 10^6 or less. This result suggests that the monuments themselves are stable to within a few centimetres over this timescale, and the apparent motions are probably attributable to the uncertainty in original observations. Possible slumping of ground (or of buildings, as at Chlemutsi), lightning strikes, and acts of war are more serious threats to monument stability in this region than is occasional shaking in earthquakes. We therefore conclude that the relative displacements, by tens of centimetres, of monuments in the more tectonically active part of the network represent (within the uncertainties) true crustal motions.

The observations reported here illustrate both the strengths and the limitations of surveys based on old triangulations. The length of the time interval involved ensures that, in this region at least, the signal is considerably greater than the noise, allowing a coarse-grained picture of the deformation to be obtained. This time interval, furthermore, is comparable with the interval (typically a few hundred years) between earthquakes on the same fault; therefore, the measurement covers a timescale that is close to that of the phenomenon of interest. The price that must be paid is that there is often no information on translation, rigid-body rotation or uniform dilatation of the network. There is also a lower limit on the scale of deformation that can be investigated: the spacing between the points in the original survey is of the order of 60 km—about twice the spacing between the active faults in the central part of the network (Fig. 1). In addition, the 1890 survey—like most classical triangulation networks—carries no precise height determination, so only horizontal relative displacements may be recovered.

The pattern of geodetic strain in central Greece seems to be one of the relative motion of a few rigid blocks, several tens to a hundred kilometres on a side, separated by regions a few tens

of kilometres wide in which strain accumulates rapidly. The displacements of PV, MV and MP (Fig. 2) on the north side of the Gulf of Korinthos are about 60 cm with respect to the Peloponnese. In turn, the four most northerly sites in the network seem to have moved with respect to the sites immediately north of the Gulf of Korinthos by about 40 cm. The two most seismically active parts of the region covered by the network are the Gulfs of Korinthos and Evvoia (Fig. 1), and these are also the most dramatic morphological features of the region, being large gulfs bounded by active normal faults. The geodetic observations are in qualitative agreement with the seismic data, in that the relative displacements within the network can largely be accounted for by movement across these two gulfs.

The only quantitative comparison that we have made between geodetic and seismic observations is in terms of the total displacement across the region. The preferred value for the displacement expressed in earthquakes (~ 70 cm, Table 1) is about 60% of the geodetically determined displacement. Such close agreement between the seismic and total strain is perhaps surprising in view of the abundant field evidence for aseismic deformation processes in near-surface rocks. It has been suggested previously^{6,23} that seismic strain may make up 50% or more of the total strain in some deforming continental belts (including the Aegean area). This study provides independent support for that suggestion from a part of one of these belts.

A final note of caution is necessary. Although the results reported here appear to identify several undeforming regions, these regions all contain faults that seem to have been active during the Quaternary (see ref. 11 for a summary). The appearance of rigidity may thus be the result of our inability to resolve a strain of less than about one part in 10^6 . It is perhaps more reasonable to regard these 'blocks' as regions that are deforming about an order of magnitude more slowly than the zones that separate them but which are, nonetheless, accumulating strain that could lead to earthquakes. It seems probable that the greater precision afforded by GPS surveys could eventually yield information about the accumulation of strain in such regions that would be valuable in earthquake hazard assessment. □

Received 10 August 1990; accepted 16 January 1991.

1. Molnar, P. *Nature* **355**, 131–137 (1988).
2. England, P. C. & Jackson, J. A. *Rev. Earth Sci.* **17**, 197–226 (1989).
3. Molnar, P. & Tapponnier, P. *Science* **189**, 419–426 (1975).
4. Molnar, P. & Deng, Q.-D. *J. geophys. Res.* **89**, 6203–6228 (1984).
5. Jackson, J. A. & McKenzie, D. P., *Geophys. J. R. astr. Soc.* **93**, 45–73 (1988).
6. Ekström, G. & England, P. C. *J. geophys. Res.* **94**, 10231–10258 (1988).
7. Lichten, S. M. & Bertiger, W. I. *Bull. Geodes.* **63**, 167–189 (1989).
8. Dixon, T. et al. *Eos* **71**, 1051–1056 (1990).
9. Hartl, H. *Comptes Rendus de la Seizième Conférence Générale de l'Association Géodésique Internationale II* (1911).
10. McKenzie, D. P. *Geophys. J. R. astr. Soc.* **30**, 109–185 (1972).
11. Ambraseys, N. N. & Jackson, J. A. *Geophys. J.* **101**, 663–708 (1990).
12. Beuttlér, G., Bauerisma, I., Gurtner, W. & Rothacher, M. in *GPS Techniques Applied to Geodesy and Surveying* 363–380 (Springer, Berlin, 1988).
13. Cross, P. A. *Proc. XIX Int. Congress Federation International de Géometrie* **5**, 336–348 (FIG, Helsinki, 1990).
14. Pope, A. J. in *NOAA Tech. Rep. N0565NGS1* (NOAA, Washington, DC, 1976).

15. Hatch, R. & Larsen, K. in *Proc. 1st Int. Symp. on Precise Positioning with GPS*, Rockville 285–297 (National Geodetic Information Center, NOAA, Rockville, Maryland, 1985).
16. Kostrov, V. V. *Izv. Acad. Sci. USSR Phys. Solid Earth* **1**, 23–44 (1974).
17. Ekström, G. & Dziewonski, A. M. *Nature* **333**, 319–323 (1988).
18. Thucydides *History of the Peloponnesian War* (transl. Warner, R.) 94 (Penguin, London 1972).
19. McKenzie, D. P. *Geophys. J. Roy. astr. Soc.* **55**, 217–254 (1978).
20. McKenzie, D. P. & Jackson, J. A. *Earth planet. Sci. Lett.* **65**, 182–202 (1983).
21. Kissel, C., Laj, C. & Mazaud, A. *Geophys. Res. Lett.* **13**, 1446–1449 (1986).
22. Kissel, C. & Laj, C. *Tectonophysics* **146**, 183–201 (1988).
23. Jackson, J. A. & McKenzie, D. P. *Geophys. J. Roy. astr. Soc.* **77**, 185–264 (1984).

ACKNOWLEDGEMENTS. This work was supported by a grant from Oxford University to P.E. and B.P., by the Hellenic National Committee on Geodesy and Geophysics, by Robert Saunders Ltd (mountain tents) and by the Royal Society. We thank K. Argiros, V. Balis, T. Bickell, S. Felekis, P. Milas, L. Moresi, R. Newman, K. Orphanos, A. Papafitsorou, P. Powell, T. Strange, Ch. Viachos and S. Watson for assistance in the field. The help of Kim Kastens and the Lamont–Doherty Geological Observatory Aegean Geodetic Project is gratefully acknowledged; we are particularly grateful to Ken Hurst for manning the Dionysos station.

Three-dimensional structure of plant light-harvesting complex determined by electron crystallography

Werner Kühlbrandt & Da Neng Wang

European Molecular Biology Laboratory, Meyerhofstrasse 1, 6900 Heidelberg, Germany

The structure of the light-harvesting chlorophyll *a/b*-protein complex, a membrane protein serving as the major antenna of solar energy in plant photosynthesis, has been determined at 6 Å resolution by electron crystallography. Within the complex, three membrane-spanning α helices and 15 chlorophyll molecules are resolved. There is an intramolecular diad relating two of the α helices and some of the chlorophylls. The spacing of the chlorophylls suggests energy transfer by delocalized exciton coupling and Förster mechanisms.

SOLAR energy is collected in photosynthetic membranes of all green plants by the light-harvesting chlorophyll *a/b*-protein complex associated with photosystem II (LHC-II). The energy is passed on to the photosynthetic reaction centres I and II where it is used to generate a potential gradient across the membrane. This provides the energy for oxygen evolution, ATP synthesis and, ultimately, carbon fixation.

LHC-II is isolated from chloroplast membranes as a trimer¹ of three identical or nearly identical monomers², each consisting of one polypeptide (233 amino acids³, relative molecular mass of 25,000) and non-covalently attached chromophores (eight chlorophyll *a*, seven chlorophyll *b* (ref. 1) and several carotenoids⁴). Altogether, LHC-II binds ~50% of all chlorophyll on earth. The complex forms large, well-ordered two-dimensional crystals⁵ (layer group p321, $a = 127$ Å, $c = 60$ Å). Henderson and Unwin developed a method for determining the structure of biological macromolecules by electron diffraction, electron microscopy and image processing of two-dimensional crystals⁶, which has recently been extended to high resolution⁷⁻⁹. By this method, the phases of structure factors are determined directly from the Fourier transform of electron micrographs to a resolution that is limited by the finest detail reproduced in the image. Untilted crystals yielded phases to 3.4 Å (ref. 5) whereas image transforms of highly tilted crystals tended to be weak at high resolution in the direction perpendicular to the tilt axis. In this study, data extending to 6 Å resolution were used (Fig. 1). Future analysis at higher resolution should be possible as more phase data from highly tilted two-dimensional crystals are incorporated.

Polypeptide structure

A three-dimensional map of LHC-II (Figs 2 and 3) was generated from 2,737 structure factors in the asymmetric unit. Figure 2 shows the trimeric complex, consisting of three monomers related by crystallographic symmetry. The overall shape of the trimer is roughly cylindrical, with a diameter of 73 Å and a thickness of 60 Å. The space around the perimeter and a central region of triangular cross-section are almost featureless and probably contain unresolved or partly disordered lipid and detergent. A part of the map of roughly elliptical, 30 Å × 50 Å cross-section which we believe to be the LHC-II monomer is shown in Fig. 3. Well resolved features highlighted in colour,

are found mainly in the 40 Å-thick central section of the map which presumably coincides with the lipid bilayer.

The most prominent features of the map are three rod-like regions of density (purple) representing three membrane-spanning α helices, labelled A, B and C in Fig. 3. The central helices A and B are inclined by 25° and 31° with respect to the membrane normal and slightly twisted round one another, with a contact distance of about 10 Å. This distance and the included angle of ~56° are characteristic for close-packed helices of this polarity in soluble proteins¹⁰. Helices A and B are ~46 Å and 49 Å long, corresponding to about 31 and 33 residues, respectively. They are thus longer than all membrane-spanning α helices of bacteriorhodopsin⁹ and of the bacterial reaction centre¹¹. The upper parts of helices A and B are related by approximate, non-crystallographic twofold symmetry. The lower third of helix A but not of helix B is noticeably kinked and almost parallel to the membrane normal. The non-crystallographic twofold symmetry does thus not extend to the lower half of the complex.

Helix C is inclined by 11° with respect to the membrane normal. Its length of about 30 Å, corresponding to 20 residues, is similar to that of the shorter transmembrane helices in the two other membrane proteins^{9,11}. This helix probably delineates the hydrophobic core of the lipid bilayer which is 30 Å thick¹². A side view (Fig. 3c) shows that helices A and B extend vertically ~7 and 9 Å beyond the top of helix C. Taking the total bilayer thickness as 40–45 Å (ref. 13) this means that they protrude somewhat beyond the membrane surface. There is no evidence of other polypeptide structures spanning the membrane.

At this contour level, little detail is visible above or below the central 40 Å section, probably due to the lack of distinct secondary structure in this region, or possibly because this part of the polypeptide is disordered. At a lower contour level (not shown), several features appear which we have tentatively assigned to loops connecting the helices and to the amino and carboxy termini of the polypeptide. On the upper side of the complex, two hook-like extensions of helices A and B with the same approximate two-fold symmetry protrude above the membrane surface, linking the two halves of the monomer which appear roughly triangular in projection (Fig. 3a). Helix C seems to be connected to the hook on helix A, whereas the hook on helix B ends in a convoluted structure which is not connected to another helix and therefore is likely to represent one of the polypeptide termini. On the other side of the membrane (Fig. 3b), helix B seems to be linked to helix C. Two roughly linear regions of density are connected to helix A which may lead to the other polypeptide terminus. The surface on this side of the membrane appears more or less flat.

Three hydrophobic membrane-spanning α helices of about 20 residues each have been predicted by the polypeptide sequence of LHC-II (refs 14, 15). The map shows that helices A and B are considerably longer than expected. The polypeptide sequence revealed an internal homology with 37% sequence identity of residues 47 to 79 and 162 to 194 in the mature protein^{16,17}. Since the homology relates equivalent segments of two predicted helices, it seems likely that these domains should be identified with the upper halves of helices A and B and their extensions which are related by twofold symmetry. The positions

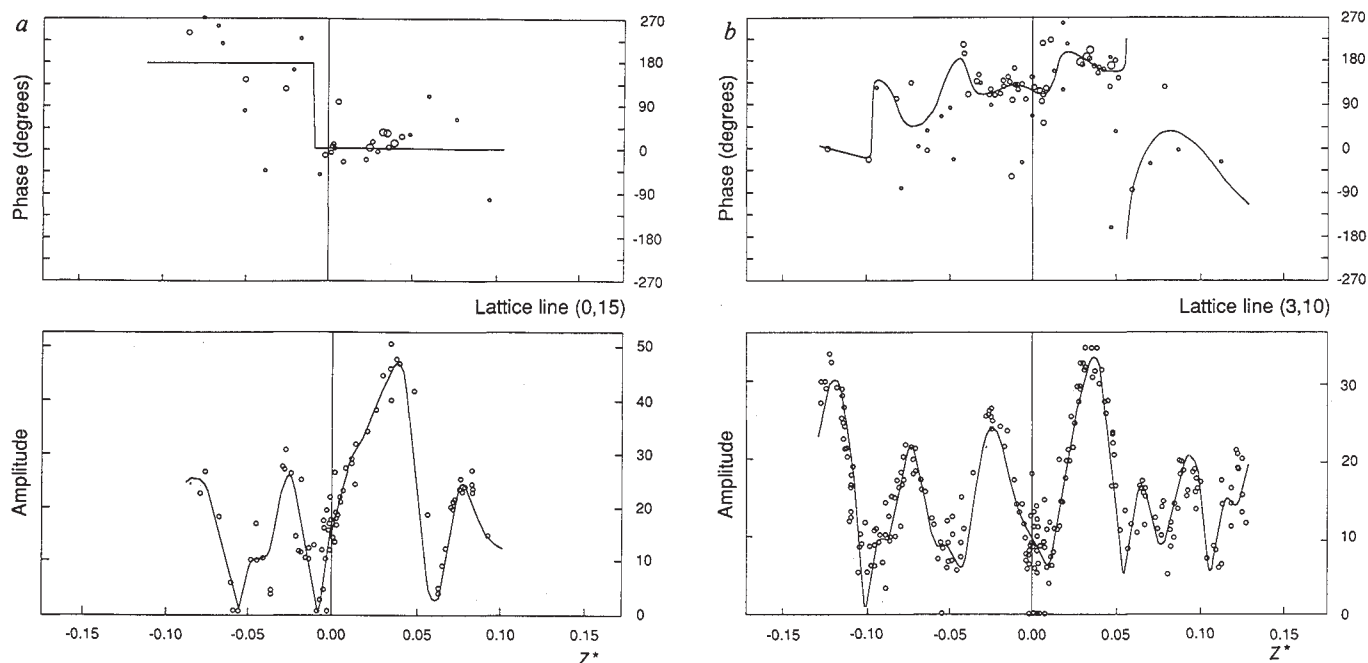


FIG. 1 Variation of phases and amplitudes of structure factors along two out of 115 lattice lines in the asymmetric unit, to 6 Å resolution. The size of circles in the upper half relates to the accuracy of phase measurement (see ref. 8), with larger circles indicating higher accuracy. Due to crystal symmetry, lattice lines with indices (0,k) have real phases of 0° or 180°. *a*, Lattice line (0,15); 7.33 Å resolution at $z^*=0$. *b*, Lattice line (3,10); 9.33 Å resolution at $z^*=0$.

METHODS. Two-dimensional crystals measuring up to 10 µm in diameter were grown from LHC-II isolated from pea leaves and specimens for electron crystallography were prepared in the presence of tannin⁵. Electron diffraction patterns yielded structure factor amplitudes to 3.2 Å resolution in all directions except for a 30° missing cone (D.N.W. and W.K., manuscript submitted). Low-resolution amplitudes to 30 Å were taken from image transforms. High-resolution images of tilted and untilted two-dimensional crystals were recorded on Agfa Scientia film in the JEOL 100B cryo-electron microscope equipped with a field emission gun³¹ at Berkeley, California, in spotscan

mode³²⁻³⁴ at a specimen temperature of -113 °C, with an electron dose of 5 to 8 e Å⁻² and magnification ×57,000. Films were developed for 12 min in full-strength Kodak D-19 developer. Areas with high crystallinity selected by optical diffraction were digitized in 6,000×6,000 7-µm steps on a Perkin-Elmer 1010 GM microdensitometer. In total, 20 out of about 1,000 images were processed, including four each recorded at tilt angles of 0°, 20° and 30°, six at 45° and two at 60°. The highest actual tilt angle was 57.2°. Image transforms were corrected for defocus gradient, lattice distortions, astigmatism and beam tilt^{8,9}. Phases of structure factors were combined, minimizing the residual phase error for each new image. Phase residuals for individual images ranged from 22° to 33° for all reflections above background. Curves were fitted³⁵ to amplitudes and phases for each lattice line and sampled at (140 Å)⁻¹ to generate a three-dimensional set of structure factors which was 84% complete. The actual resolution in z direction was estimated at about 8 Å from the point spread function for the 30° missing cone³⁶.

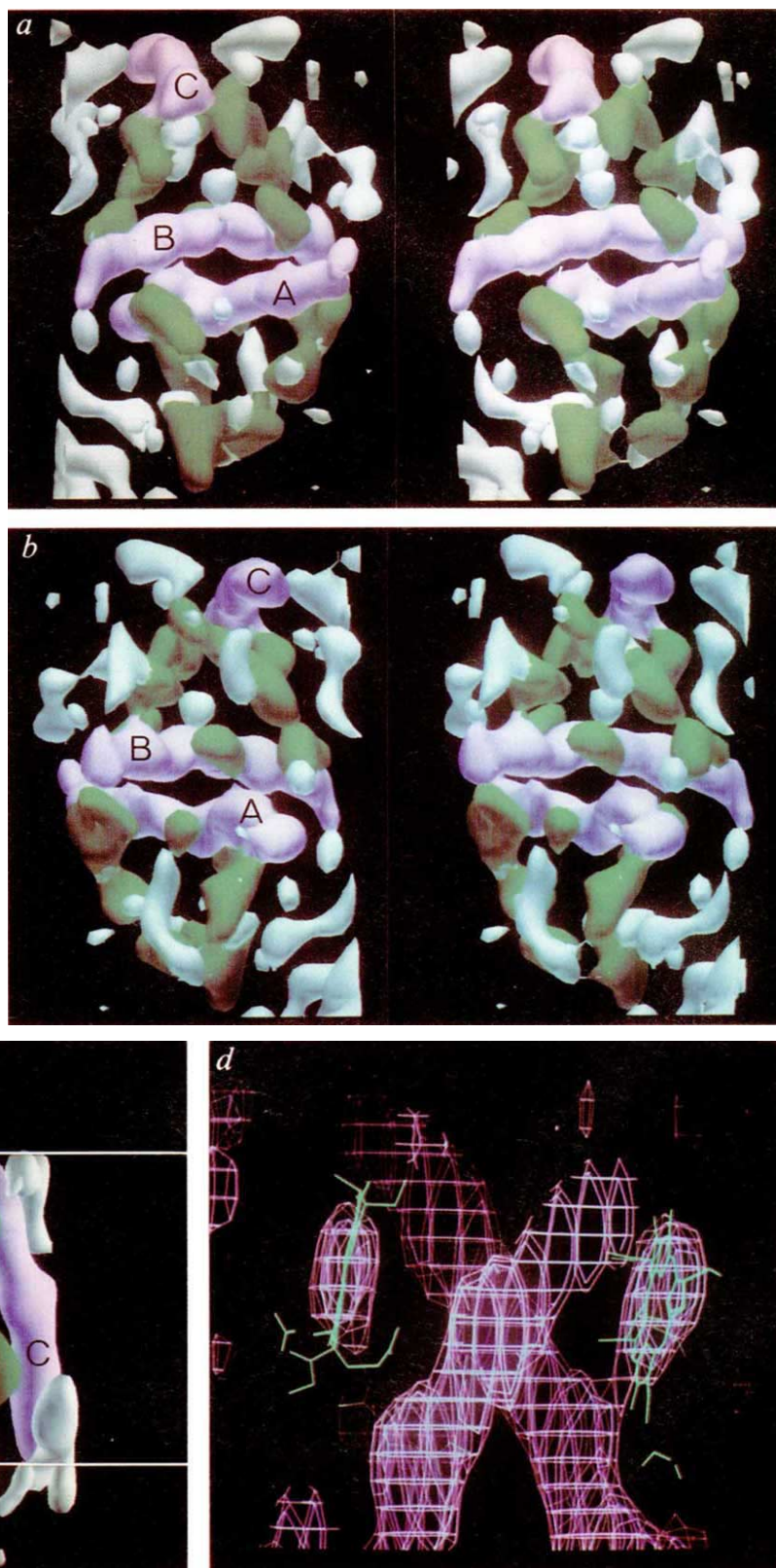


FIG. 2 Three-dimensional map of the LHC-II trimer at 6 Å resolution seen from a position above the membrane plane. Well resolved rod-like features (purple) represent α helices. Small, roughly ellipsoidal centres of density (green) were assigned to porphyrin head groups of chlorophyll molecules. Partly resolved regions of the polypeptide at the membrane surface are white. The trimer consists of three apparently identical monomers, related by crystallographic threefold symmetry. Featureless regions in the centre and around the perimeter of the trimer probably contain disordered lipid and detergent. Solvent flattening³⁷ reduced the phase error only by about 10°, reflecting the high quality of phases. The map was generated without solvent flattening by standard crystallographic procedures and coloured using a program written by S. Fuller.

of the homologous domains in the sequence suggest that they are on the same surface as the N terminus^{14,15} which is thought to be located on the outer, stromal surface¹⁸. By implication, the map in Figs 2 and 3a is viewed from the stromal side. We therefore propose a topography of the polypeptide as shown in Fig. 4, even though we cannot trace the amino-acid chain with certainty at this resolution.

LHC-II, a eukaryotic membrane protein, bears no resemblance to the three prokaryotic membrane proteins, the structures of which have been determined at high resolution: bacteriorhodopsin⁹, the bacterial reaction centre¹¹, and matrix porin from *Rhodobacter capsulatus*¹⁹. However, LHC-I, the eukaryotic antenna complex closely associated with photosystem I in chloroplasts, should resemble LHC-II structurally. The

FIG. 3 Views of the LHC-II monomer. Colouring of α helices and chlorophyll porphyrins in *a* to *c* as in Fig. 2. Polypeptide at the membrane surface and density in the membrane belonging to the adjacent monomer is white. Two stereo pairs show the complex as seen from above (*a*) and below (*b*) the membrane. The side view (*c*) is from a central position in the lipid bilayer outside the trimer. Well resolved features are largely confined to the level of the hydrophobic part of the bilayer. The monomer contains three membrane-spanning α helices, labelled A, B and C, plus 15 chlorophyll molecules, labelled 1 to 15 in *c* (chlorophylls 6 and 12 are hidden behind helix B). Helices A and B are inclined by about 30° with respect to the membrane normal. Helix C runs approximately perpendicular to the membrane plane. The upper parts of helices A and B and the four chlorophylls nearby (numbers 1 and 5, 2 and 6) are related by local, twofold symmetry (*a*). This symmetry is not evident on the reverse side of the complex (*b*). Helices A and B seem to protrude at the upper membrane surface (*c*). Helix C appears to be fully embedded in the hydrophobic lipid bilayer. Chlorophyll porphyrins are arranged on two levels, corresponding roughly to the two bilayer leaflets. Some of the chlorophylls have protrusions which may be part of the phytol chains. *d*, Detail of helices A and B and of chlorophylls 1 and 5, related by the non-crystallographic diad. A molecular model of chlorophyll *a* porphyrin was adapted from bacteriochlorophyll *a* coordinates²⁰ and fitted to 15 centres of density in the map with similar dimensions.



polypeptide sequence of both complexes indicates significant sequence identity¹⁶, the most highly conserved regions being those which are related by internal homology in LHC-II.

Arrangement of chlorophylls

Another striking feature of the map is the presence of several small, distinct regions of density (green in Figs 2 and 3), surrounding the central pair of α helices in the hydrophobic interior of the complex. These are clearly not part of the polypeptide since this is forced into an α -helical or β -sheet conformation by the hydrophobic environment. Both types of secondary structure can be distinguished easily as such at 6 Å resolution. Most of these centres of density have an ellipsoidal, somewhat angular shape, measuring about 8 Å in diameter and about 4 Å in thickness which leaves no doubt that they represent the planar porphyrin ring systems of chlorophyll molecules. We assigned chlorophylls to 15 such centres of density, in agreement with biochemical findings which indicated a chlorophyll: polypeptide molar ratio of 15:1 (ref. 1). As expected, chlorophyll *a* and *b* cannot be distinguished at this resolution. Figure 3*d* shows how chlorophyll porphyrin rings might fit into two of these regions. All assigned chlorophyll densities were roughly of this size except number 13 which was smaller and may prove not to be a porphyrin at higher resolution. A feature extending into the hydrophobic bilayer near the bottom of helix C has the correct shape and size for a porphyrin but seems too distant from any resolved polypeptide. Otherwise, only few small islands of density are visible in the interior of the complex. These may belong to the linear hydrocarbons of phytol chains or carotenoids which are not resolved.

The side view (Fig. 3*c*) reveals that the chlorophyll head groups are arranged on two levels, corresponding roughly to the upper and lower leaflet of the lipid bilayer. The upper level (Figs 3*a* and 5*a*) contains eight chlorophylls (numbers 1 to 8), at least four of which (numbers 1 and 5, 2 and 6) are related by the same local twofold symmetry that was found for the upper parts of helices A and B, indicating that there are conserved chlorophyll binding sites in the homologous sequence. Figure 3*d* illustrates the local non-crystallographic symmetry of helices A and B and of chlorophylls 1 and 5. The seven lower chlorophylls (numbers 9 to 15), like the polypeptide on this side of the complex, are apparently not related by twofold symmetry (Figs 3*b* and 5*b*).

Chlorophyll porphyrins are oriented roughly perpendicular to the membrane plane, at an average angle of about 10° from the membrane normal. The centre-to-centre distances between nearest neighbours on each level fall into a narrow range of 9 to 14 Å. Chlorophylls on the upper level are arranged in a ring-like fashion (Fig. 5*a*), so that each has two neighbours. The plane of this ring is inclined to bring chlorophylls 6, 7 and 8 closer to the lower level. The lower chlorophylls describe an approximate crescent shape (Fig. 5*b*) which slopes upwards, with chlorophylls 9, 10 and 11 closest to chlorophylls 6, 7 and 8 on the upper level, at a distance of 13 to 14 Å.

The LHC-II polypeptide contains only three histidines which is the usual ligand for Mg^{2+} of chlorophylls^{11,20}. Most chromophores in the complex must therefore be bound by other residues or in other ways. Several chlorophylls seem to be suspended between the polypeptide on the membrane surface and the membrane-spanning α helices. The centres of these chlorophyll porphyrins seem rather too distant from the polypeptide for direct coordination of amino-acid side chains to Mg^{2+} , suggesting that water molecules intervene, or that these chlorophylls are held in place by hydrogen bonding to the perimeter of the porphyrin rings.

Compared with LHC-II, the relationship between chromophores and polypeptide is reversed in the two bacteriochlorophyll-protein complexes of known structure, where pigment molecules are enclosed in a polypeptide cage. The light-harvesting bacteriochlorophyll *a*-protein complex from *Pro-*

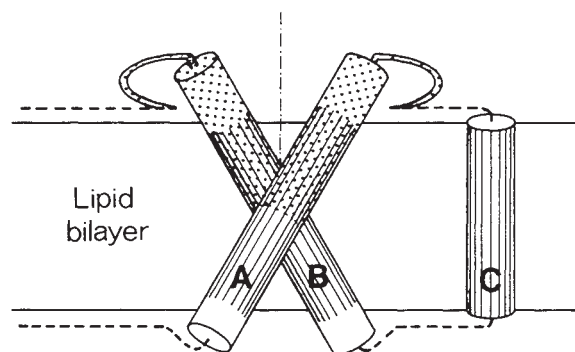


FIG. 4 Proposed topography of the LHC-II polypeptide in the photosynthetic membrane. Hydrophobic regions of helices A, B and C are striped. Shaded regions of helices A and B and their hook-shaped extensions at the top of the complex are related by a non-crystallographic twofold axis running roughly perpendicular to the membrane plane (dash-dotted line) and by internal polypeptide homology. Dashed lines represent domains of the polypeptide on the membrane surfaces that are less well resolved in the map. The order of the three helices in the polypeptide sequence seems to be B, C and A, beginning at the N terminus. The homologous regions comprise the upper half to two thirds of helices A and B and their hook-like extensions. These regions are on the same side as the N terminus^{14,15} which is thought to be located on the outer, stromal side of the membrane¹⁸.

*thecochloris aestuarii*²⁰ is a soluble protein, and its polypeptide forms a β barrel to shield the hydrophobic chromophores in the interior. By contrast, the chlorophyll-containing region of LHC-II is surrounded by lipid and need not be shielded. In the bacterial reaction centre¹¹, the prosthetic groups are partly surrounded by a total of ten α helices creating a special environment which must be responsible for the large red shift of the absorption maximum. In plant photosynthesis, the red shift is less pronounced and therefore such a special protein environment is not required in LHC-II. The role of the polypeptide, and in particular of the tilted helices A and B, seems to be primarily that of a scaffold for packing the chromophores into a small volume, ensuring optimal spacing and orientation for energy transfer.

The trimer

Trimer formation appears to be mediated mainly by interaction between monomers at the membrane surface, notably at the upper ends of helices B which are less than 7 Å apart, because there are few polypeptide contacts in the hydrophobic interior. A region of low density was chosen as the molecular boundary (dashed line in Fig. 2), so that within one monomer, the same polypeptide chain was closest to all chlorophylls. All other possible boundaries place several chlorophylls close to polypeptides of two different monomers. A boundary between helices A and B could be proposed on the basis that other membrane proteins interact through contacts between hydrophobic α helices^{9,11}. Gene duplication, which probably gave rise to the local twofold axis in LHC-II, tends to result in intramolecular rather than intermolecular symmetry, and therefore such a boundary appears unlikely. Nevertheless, the molecular boundary shown here should be regarded as preliminary.

There are several reasons to believe that the trimer is the native, functional conformation of LHC-II. Analytical ultracentrifugation has shown that the trimeric complex is highly stable in detergent solution¹. Fluorescence kinetics of isolated LHC-II indicated that rapid energy transfer from chlorophylls *b* to *a*, which is characteristic for the functional complex, breaks down as the trimer dissociates²¹. Freeze-fracture electron microscopy of native thylakoid grana revealed a large number of 80 Å particles²² of the same shape and size as on fracture faces of two-dimensional crystals of LHC-II (ref. 23). Finally, most other known light-harvesting complexes have threefold symmetry^{20,24-26}. Energy transfer from an antenna of this sym-

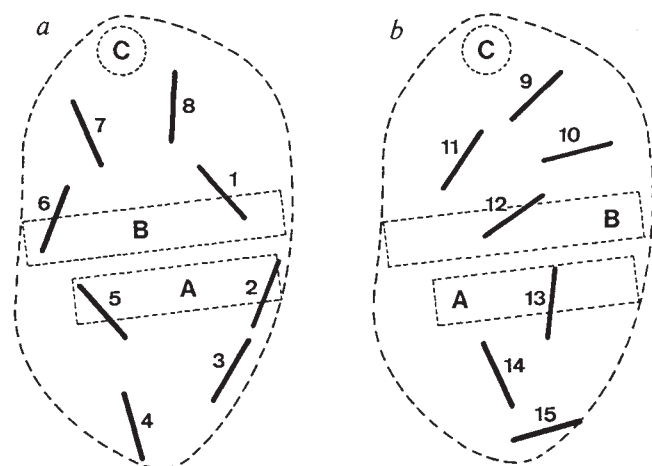


FIG. 5 Approximate position of chlorophylls on the upper level (a) and on the lower level (b) in the complex. Chlorophyll porphyrins, represented by black bars, are oriented roughly perpendicular to the membrane plane. The outlines of the monomer and of helices A, B and C are dashed. Note that a and b are seen from above the membrane and thus b appears mirror-symmetric to Fig. 3b. Chlorophyll 14 is attached to helix A, chlorophyll 10 to helix B and chlorophylls 7 and 8 to helix C. Two chlorophylls (numbers 10 and 15) are only partly resolved from the non-helical polypeptide at the membrane surface. A preliminary estimate of energy transfer rates between chlorophylls was obtained from expressions²⁹ for the interaction energy and energy transfer times for the two cases of delocalized exciton coupling and Förster transfer, assuming that all chlorophylls were chlorophyll a. The dipole strength of chlorophyll a was taken as 16.8 Debye (ref. 38), the natural fluorescence decay time as 1.5×10^{-8} s (ref. 39) and the critical distance for Förster transfer at orientation factor 1 as 90 Å (ref. 29). Because the orientation of porphyrin rings cannot be determined with certainty at present, the factor for random orientation, 0.66 (ref. 29), was used in the calculation.

metry to other molecules in the same plane is less dependent on the rotational alignment of the molecules. We therefore conclude that Fig. 2 shows the structure of the complex as it exists in the photosynthetic membrane.

The chlorophyll concentration in the whole volume of the LHC-II trimer is 0.3 M. The chromophore concentration in the trimer of the soluble, bacterial complex from *P. aestuarii*²⁰ is 0.11 M, even though the local concentration in the central hydrophobic core (0.94 M) is slightly higher than in the corresponding region of LHC-II (0.7 M). Moreover, LHC-II uses only 15 amino-acid residues per bound chlorophyll, whereas the soluble

antenna complex uses 50. Overall, the plant light-harvesting complex is thus about three times more effective at organizing chlorophylls and, hence, as a molecular antenna of solar energy.

Energy transfer

In photosynthesis, energy is transferred between chlorophyll molecules mainly by two mechanisms, delocalized exciton coupling and Förster transfer²⁷. Exciton coupling occurs between similar or identical chromophores at distances up to 20 Å, depending on the angle between them. Energy oscillates between donor and acceptor without being localized on either molecule. At distances greater than 20 Å, Förster transfer²⁸ is more favourable. This is a longer-range dipole-dipole interaction through space whereby the acceptor molecule is excited as the donor molecule drops to the ground state.

Based on the observed chlorophyll positions we estimated the rate of energy transfer between chlorophylls in the monomer and in the trimer, assuming that all chlorophylls were chlorophyll a. Our calculation indicated that delocalized exciton coupling occurs between chlorophylls in one LHC-II monomer within about 10^{-13} to 10^{-14} s, even when taking the orientation of Qy transition dipole moments as random. Optimal orientation would result in a transfer rate that is about 2 to 3 times faster. Since every chlorophyll can absorb, on average, only one photon per second even in strong sunlight²⁹, delocalized exciton-coupling seems essential for efficient light-harvesting. Energy transfer between monomers in the same trimer is more likely to proceed by the Förster mechanism because of the greater distance of chromophores.

Owing to interspersed lipid, the average shortest distance between chlorophylls in adjacent trimers in the thylakoid membrane is probably much longer than 10–15 Å, the minimum value assuming van der Waals' contact. This means that Förster transfer is the likely mechanism of energy transmission over long distances in the membrane. The orientation of chlorophylls 2, 3 and 15 at the outer hydrophobic perimeter of the trimer may be favourable for energy transmission to neighbouring light-harvesting complexes or reaction centres.

The site of primary charge separation, to which energy must ultimately be transferred, is the special pair of chlorophylls in the photosystem II reaction centre complex. By homology with the bacterial reaction centre³⁰, it is located on the inner, luminal side of the membrane. If the inferred orientation of LHC-II in the photosynthetic membrane is correct, the special pair is on the same side of the membrane as the seven lower chlorophylls. □

Received 28 December 1990; accepted 15 February 1991.

- Butler, P. J. G. & Kühlbrandt, W. *Proc. natn. Acad. Sci. U.S.A.* **85**, 3797–3801 (1988).
- Kühlbrandt, W. *Nature* **307**, 478–480 (1984).
- Cashmore, A. R. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2960–2964 (1984).
- Siefermann-Harms, D. *Biochim. biophys. Acta* **811**, 325–355 (1985).
- Wang, D. N. & Kühlbrandt, W. *J. molec. Biol.*, in the press (1991).
- Henderson, R. & Unwin, P. N. T. *Nature* **257**, 28–32 (1975).
- Baldwin, J. M. & Henderson, R. *Ultramicroscopy* **14**, 319–336 (1984).
- Henderson, R., Baldwin, J. M., Downing, K. H., Lepault, J. & Zemlin, F. *Ultramicroscopy* **19**, 147–178 (1986).
- Henderson, R. *et al. J. molec. Biol.* **213**, 899–929 (1990).
- Chothia, C. A. *Rev. Biochem.* **53**, 537–572 (1985).
- Deisenhofer, J., Epp, O., Miki, K., Huber, R. & Michel, H. *Nature* **318**, 618–624 (1985).
- Deisenhofer, J. & Michel, H. *Science* **245**, 1463–1473 (1989).
- Yeates, T. O., Komyia, H., Rees, D. C., Allen, J. P. & Feher, G. *Proc. natn. Acad. Sci. U.S.A.* **84**, 6438–6442 (1987).
- Karlin-Neumann, G. A., Kohorn, B. D., Thornber, J. P. & Tobin, E. M. *J. molec. Appl. Gen.* **3**, 45–61 (1985).
- Bürgi, R., Suter, F. & Zuber, H. *Biochim. biophys. Acta* **890**, 346–351 (1987).
- Hoffmann, N. E. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 8844–8848 (1986).
- Kühlbrandt, W. & Downing, K. H. *J. molec. Biol.* **207**, 823–828 (1989).
- Mullett, J. E. *J. biol. Chem.* **258**, 9941–9948 (1983).
- Weiss, M. S., Wacker, T., Weckesser, J., Welte, W. & Schulz, G. E. *FEBS Lett.* **267**, 268–272 (1990).
- Matthews, B. W., Fenna, R. E., Bolognesi, M. C., Schmid, M. F. & Olson, J. M. *J. molec. Biol.* **131**, 259–285 (1979).
- Ide, J. P., Klug, D., Kühlbrandt, W., Giorgi, L. B. & Porter, G. *Biochim. biophys. Acta* **893**, 349–364 (1987).

- Simpson, D. J. *Carlsberg Res. Comm.* **44**, 305–336 (1979).
- Kühlbrandt, W., Thaler, T. & Wehrli, E. *J. Cell Biol.* **96**, 1414–1424 (1983).
- Schirmer, T., Bode, W., Huber, R., Sidler, W. & Zuber, H. *J. molec. Biol.* **184**, 257–277 (1985).
- Stark, W., Kühlbrandt, W., Wildhaber, I., Wehrli, E. & Mühlethaler, K. *EMBO J.* **3**, 777–783 (1984).
- Ford, R. C., Hefti, A. & Engel, A. *EMBO J.* **9**, 3067–3075 (1990).
- Sauer, K. in *Encyclopaedia of Plant Physiology Vol 19: Photosynthesis III* (eds Staehelin, L. A. & Arntzen, C. J.) 85–97 (Springer, Berlin, 1986).
- Förster, Th. in *Modern Quantum Chemistry Part 3* (ed. Sinanoglu, O.) 93–137 (Academic Press, London, 1965).
- van Grondelle, R. *Biochim. biophys. Acta* **811**, 147–195 (1985).
- Michel, H. & Deisenhofer, J. in *Encyclopaedia of Plant Physiology Vol 19: Photosynthesis III* (eds Staehelin, L. A., Arntzen, C. J.) 371–387 (Springer, Berlin, 1986).
- Hayward, S. B. & Glaeser, R. M. *Ultramicroscopy* **5**, 3–8 (1980).
- Henderson, R. & Glaeser, R. M. *Ultramicroscopy* **16**, 139–150 (1985).
- Downing, K. & Glaeser, R. M. *Ultramicroscopy* **20**, 269–278 (1986).
- Bullough, P. & Henderson, R. *Ultramicroscopy* **21**, 223–230 (1987).
- Agard, D. A. *J. molec. Biol.* **167**, 849–852 (1983).
- Glaeser, R. M., Tong, L. & Kim, S.-H. *Ultramicroscopy* **27**, 307–318 (1989).
- Wang, B. C. *Meth. Enzym.* **115**, 90–112 (1985).
- Shipman, L. L., Cotton, T. M., Norris, J. R. & Katz, J. J. *J. Am. chem. Soc.* **98**, 8222–8230 (1976).
- Brody, S. S. *Rev. scient. Instrum.* **28**, 1021–1026 (1957).

ACKNOWLEDGEMENTS. We thank R. Henderson for image processing programs, advice and encouragement, K. Dörr for preparing the complex and A. Lahm for advice on crystallographic programs. Special thanks go to K. H. Downing for the use of the cryo-electron microscope at the Lawrence Berkeley Laboratory. Financial support from the Deutsche Forschungsgemeinschaft is gratefully acknowledged.

Iron distribution in the intracluster gas of the Virgo cluster of galaxies

Katsuji Koyama, Shiro Takano & Yuzuru Tawara

Department of Astrophysics, School of Science,
Nagoya University Furo-cho, Chikusa-ku, Nagoya 464-01, Japan

DIFFUSE X-ray emission, particularly fluorescence from the iron K line, from hot gas in the clusters of galaxies provides valuable information on the structure and evolution of the clusters¹. Because heavy elements such as iron are made in stars, the presence of iron distinguishes gas that originated in galaxies from primordial gas that has remained unprocessed since formation of the cluster. Previous observations of about 20 clusters indicate that iron abundance in the hot gas is about half the cosmic value², but these observations were made with non-imaging devices which gave no information on the spatial distribution of the iron. Lack of spatial distribution can also skew the estimated mean iron abundance determined in this way. Here we report evidence of an iron abundance gradient in hot gas in the Virgo cluster of galaxies. Iron is centrally concentrated, indicating that gas at the cluster core has been enriched by galactic mass loss.

The Virgo cluster is the nearest rich cluster of galaxies and has a large optical extension of $\sim 12^\circ$ along the major axis of the cluster. It is classified as an irregular cluster in which the distribution of the galaxies is unrelaxed. The galaxy distribution has two major concentrations of galaxies around M87 and M49 (ref. 3). We have conducted X-ray scanning observations from south to north along the major axis, which runs over M49 and M87, using the Large Area Counter (LAC)⁴ onboard the Ginga satellite⁵. The beam size of the LAC is $\sim 1^\circ$ and 2° (full width at half maximum, FWHM) along and perpendicular to the scan path, respectively. The LAC is sensitive in the energy band 1–37 keV with a maximum effective area of 4,000 cm². With the high sensitivity of the LAC, we found large-scale extended X-ray emission of $\sim 12^\circ$ covering the whole Virgo cluster⁶.

Figure 1a, b shows typical X-ray spectra summed for the data from the central region (within position angle 0.5° of M87) and outer region ($2\text{--}3.5^\circ$ north of M87) of the Virgo cluster, respectively. The X-ray spectrum near the central region cannot be fitted by a single-temperature model but can be fitted by a two-component model of thermal bremsstrahlung (and iron line) and power-law emission. The best-fit photon index and temperature are 1.90 ± 0.15 and 2.19 ± 0.04 keV, respectively (hereafter, we quote errors at the 90% confidence level of the parameter of interest). The best-fit models are given by the histograms in Fig. 1a.

To separate the power-law component from the optically thin thermal intracluster emission, we obtained X-ray spectra at each position of the scan path. To these spectra, we fitted the two-component model, fixing the photon index at 1.9. The free parameters of these fits are the normalization factor of the power-law component, the temperature and normalization factor of the thermal-bremsstrahlung component, and the central energy and equivalent width of iron line. The fits were generally acceptable, with the reduced χ^2 ranging from 0.6 to 1.3. The best-fit normalization factor of the power-law component becomes negligibly small when the scan angle is more than $\sim 1^\circ$ from M87. In this region we tried fits to a single-temperature thermal-bremsstrahlung model and found the fits to be acceptable with the reduced χ^2 ranging from 0.8 to 1.2. The histogram of Fig. 1b is an example of the best fit to a single-temperature thermal-bremsstrahlung model with temperature of 2.05 ± 0.16 keV.

Using additional scans near the cluster centre, we found that the power-law emission is due to the active galactic nuclei of

M87 and Seyfert 2 galaxy NGC4388 (ref. 7). Hanson *et al.*⁹ also found no hard-X-ray emission other than that from M87 and NGC4388 within $\sim 4^\circ$ of M87. We therefore conclude that the power-law component is due only to M87 and NGC4388.

Figure 2 shows the best-fit temperature distribution along the scan path. In the region within $\sim 1^\circ$ of M87, the best-fit temperature is the result of two-component fit, whereas that of the region farther than $\sim 1^\circ$ from M87 is the best-fit temperature of a single-component model. We find that the temperature of intracluster gas is not isothermal but shows significant increase from 2 keV to 3.6 keV toward the south part of the cluster.

At most scan positions, we found significant iron-line emission at ~ 6.7 keV, a value consistent with that expected from plasma at a temperature of 2.0–3.6 keV in ionization equilibrium. The active galactic nuclei M87 and NGC4388 can also be iron-line emitters. The iron equivalent width of a Seyfert 2 galaxy is generally larger than that of the other class of active galactic nuclei⁸ but the hydrogen column density of NGC4388 is $\sim 2 \times 10^{23}$ H cm² (ref. 9). In this range the equivalent width is about 200–300 eV (ref. 8). We made a rather conservative estimate of the possible contamination by the iron-line flux from M87 and NGC4388, assuming that the iron-line equivalent widths are ~ 300 eV, and found that the contamination from these active galactic nuclei is only 10% of the total iron-line flux. We therefore neglect contamination by the iron-line flux from these active galactic nuclei in the following discussion.

We derived the equivalent width of the iron line by dividing the best-fit iron flux by the flux of thermal-bremsstrahlung component at 6.7 keV. The continuum emission of the power-law component due to M87 and NGC4388 has therefore already been removed. The equivalent width can be converted to iron abundance using the calculated equivalent width of the cosmic abundance under the assumption of ionization equilibrium. The iron abundance distribution is plotted in Fig. 3 as a function of

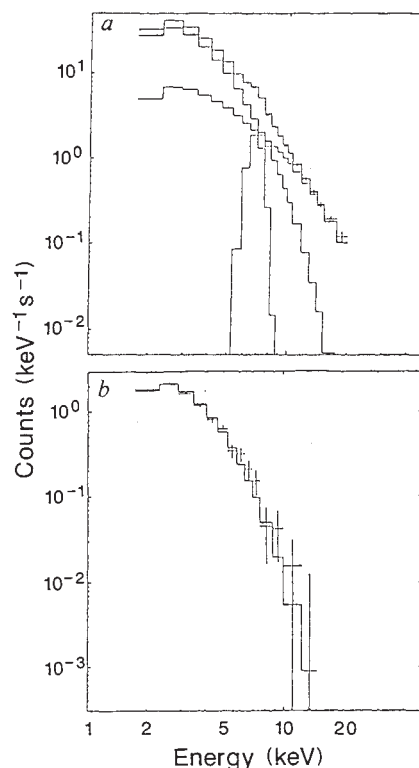


FIG. 1 X-ray spectra summed for the scan data a, within the position angle 0.5° from M87 and b, $2\text{--}3.5^\circ$ north from M87. The solid histograms are from the best-fit model described in the text.

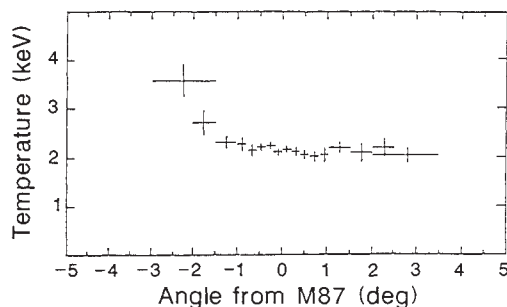


FIG. 2 The best-fit temperature distribution along the scan path from south (negative angles) to north (positive angles). In the region within $\sim 1^\circ$ of M87 the results are those of a two-component fit, and those in the outer region are the best-fit parameters of a single-temperature model (see text). The scan profile is not deconvolved from the collimator response (1° FWHM).

scan angle. The calculated equivalent width of the cosmic abundance is in the range 1.4–1.7 keV in the relevant temperature range of 2.0–3.6 keV (ref. 10). Therefore any systematic error due to a temperature gradient is at most $\sim 10\%$.

The distribution of iron abundance shows a clear peak near the cluster centre at M87. The iron abundance in the outer region of the cluster (more than 1° from M87) is about 0.1–0.2 of the cosmic value. Because the angular profile of the iron abundance is comparable to that of the collimator response (1° FWHM), we cannot resolve the detailed structure of the iron abundance within 1° of M87, but we can safely conclude that the mean iron abundance in the central region (within 1° of M87) is ~ 0.5 of the cosmic value. The previous result¹¹ of mean abundance is roughly consistent with our result for the central region. This is reasonable because the mean abundance is a luminosity-weighted mean. Because luminosity is a function of the square of gas density, the cluster centre with high gas density gives an over-estimation of the mean abundance. The gas-density-weighted mean would give a more realistic abundance. From the two-dimensional mapping of the cluster gas, we estimate that the total mass of gas in the Virgo cluster is about 10 times larger than that of the central region (within $50'$ of M87)¹². Therefore the 'true' mean abundance of the Virgo cluster should be similar to that of the outer region (0.1–0.2 of the cosmic value). As the mean abundance of the Virgo cluster is only 0.1–0.2 of the cosmic value, we conclude that the intracluster gas of Virgo is not mainly due to matter outgassed from the member galaxies. By contrast, the central region of the cluster has been contaminated by more efficient mass loss from galaxies because the galaxy distribution is denser than that in the outer region of the cluster. We note that Ponman *et al.*¹³ reported a

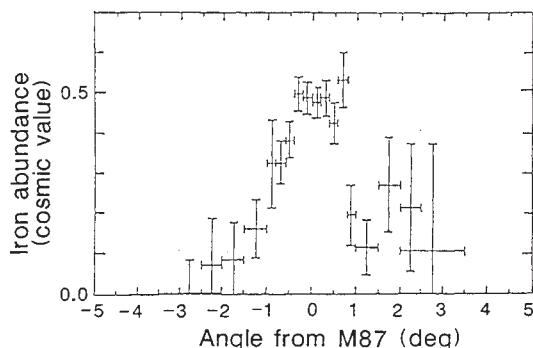


FIG. 3 Iron abundance as a function of position angle from M87, where the scan profile is not deconvolved from the collimator response (1° FWHM).

similar central concentration of iron abundance from the Perseus cluster of galaxies. Our results for the Virgo cluster, as well as those for the Perseus cluster, clearly demonstrate that spatially resolved iron-line spectroscopy is necessary to investigate the origin and evolution of the intracluster gas. $\square$

Received 23 July 1990; accepted 17 January 1991.

1. Sarazin, C. L. *Rev. mod. Phys.* **58**, 1–116 (1986).
2. Rothenflug, R. & Arnaud, M. *Astr. Astrophys.* **144**, 431–442 (1985).
3. Binggeli, B., Tamman, G. A. & Sandage, A. *Astr. J.* **94**, 251–277 (1987).
4. Turner, M. J. L. *et al. Publ. astr. Soc. Japan* **41**, 345–372 (1989).
5. Makino, F. *et al. Astr. Lett. Commun.* **25**, 223–233 (1987).
6. Takano, S. *et al. Nature* **340**, 289–290 (1989).
7. Takano, S. & Koyama, K. *Publ. astr. Soc. Japan* (in the press).
8. Awaki, H. *et al. Publ. astr. Soc. Japan* (in the press).
9. Hanson, C. G. *et al. Astr. Astrophys.* **242**, 262–266 (1990).
10. Mewe, R. *et al. Astr. Astrophys. Suppl.* **65**, 511 (1986).
11. Serlemitsos, P. J. *et al. Astrophys. J.* **211**, L63–L67 (1977).
12. Fabricant, D., Lecar, M. & Gorenstein, P. *Astrophys. J.* **241**, 552–560 (1980).
13. Ponman, T. J. *et al. Nature* **347**, 450–452 (1990).

ACKNOWLEDGEMENTS. We thank C. Day and M. Lim for comments. This work was supported by the Japanese Ministry of Education, Science and Culture.

Irradiation-driven mass transfer in low-mass X-ray binaries

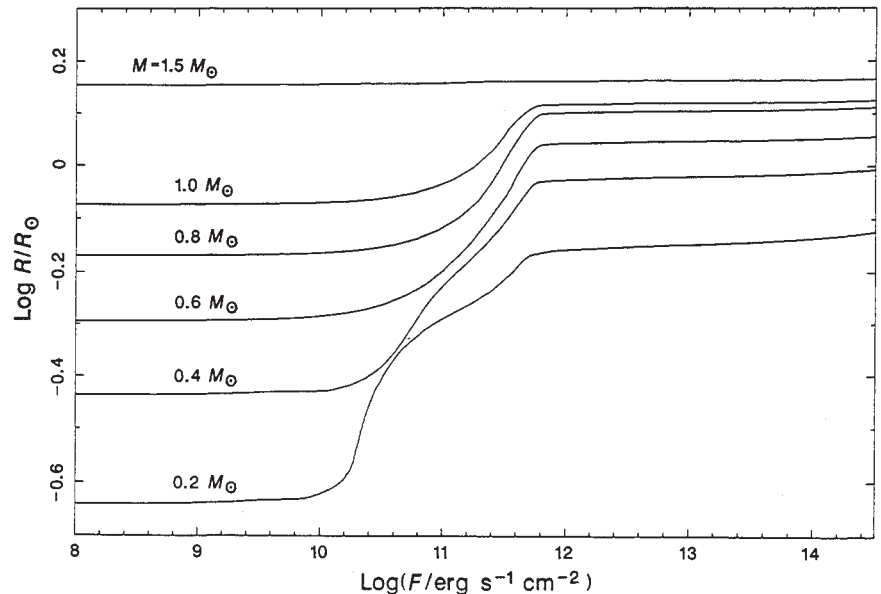
Ph. Podsiadlowski

Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

IN a low-mass X-ray binary (LMXB), mass accreted onto a neutron star from a larger companion generates a flux of X-rays which irradiate the companion. Previous studies of irradiated stars have been restricted to the effects on their outermost layers^{1–5}, or did not address conditions found in interacting binaries⁶. Here I show that, under the radiative flux typical of a LMXB, the companion will expand towards a new state of thermal equilibrium, and that this expansion provides a new mechanism to drive mass transfer onto the neutron star. The evolution of LMXBs can be drastically altered in this way. In particular, this mechanism may introduce a new evolutionary phase during which the orbital period increases, leading to larger orbital periods during and at the end of mass transfer, and (if the companion is a subgiant) shortening the duration of the LMXB phase. These modifications may help to resolve a number of observational puzzles concerning the numbers and properties of LMXBs and related objects.

The main effect of external irradiation is to change the star's effective surface boundary condition, particularly by altering the degree of ionization of hydrogen at the bottom of the irradiated layer (which defines the effective surface boundary). Figure 1 illustrates the effect of irradiation on the radius, R , for various unevolved stars of mass M as a function of external flux, F . In these calculations, the star was assumed to be in thermal equilibrium, and all the external flux was deposited below the photosphere, where the total energy deposited as a function of Thompson optical depth, τ_{Th} , was chosen to be proportional to $(1 - e^{-\tau_{\text{Th}}})$. Figure 1 shows that, for a given mass ($M \leq 1 M_\odot$), there are two flux regimes, separated by a characteristic flux $F_{\text{Ch}} \approx 10^{11} - 10^{12} \text{ erg s}^{-1} \text{ cm}^{-2}$, within which the equilibrium radius, and thus the structure of the star as a whole, changes only slowly with changing external flux. The difference between the two regimes is that, for $F \ll F_{\text{Ch}}$, hydrogen at the bottom of the irradiated layer is mainly non-ionized, whereas for $F \gg F_{\text{Ch}}$ it is completely ionized and the star is completely radiative). We find that the overall behaviour of irradiated stars is not very sensitive to how energy is deposited in its outer layer. Indeed, our results would be qualitatively very similar if we had

FIG. 1 Equilibrium radius, R , as a function of external irradiation flux, F , for unevolved stars of mass M .



deposited no energy below the photosphere but only changed the surface boundary condition (see ref. 6).

Although these calculations assume that the irradiated star is in thermal equilibrium, this will not be true initially for the irradiated secondaries in LMXBs, because we expect the timescale for the turn-on of the system's X-ray luminosity generally to be short compared with the thermal timescale of the secondary's envelope. This means that, immediately after turn-on of the X-ray flux, the secondary, which 'suddenly' has to adjust to a new surface boundary condition, will be out of thermal equilibrium and will start to relax towards its new equilibrium configuration (shown in Fig. 1). The timescale for this relaxation, t_{rel} , is approximately the thermal timescale of the convective envelope, because a radiative core, if it exists, is little affected by a change of the surface boundary conditions; that is, $t_{\text{rel}} \approx GM_2 \Delta M / 2RL$, where M_2 is the total mass of the secondary, ΔM is the mass in its convective envelope, R is its radius, and L is the secondary's internal luminosity. For $M_2 \approx 0.8 M_\odot$, for example, $t_{\text{rel}} \approx 10^7$ yr. It is worth emphasizing that the energy required to inflate the star to its new equilibrium radius is not supplied by the external radiation flux (the star is not heated from the outside), but consists almost entirely of energy generated in the secondary's core. This explains why t_{rel} is independent of the external flux. A detailed description of the relaxation process will be published elsewhere (Ph.P., manuscript in preparation).

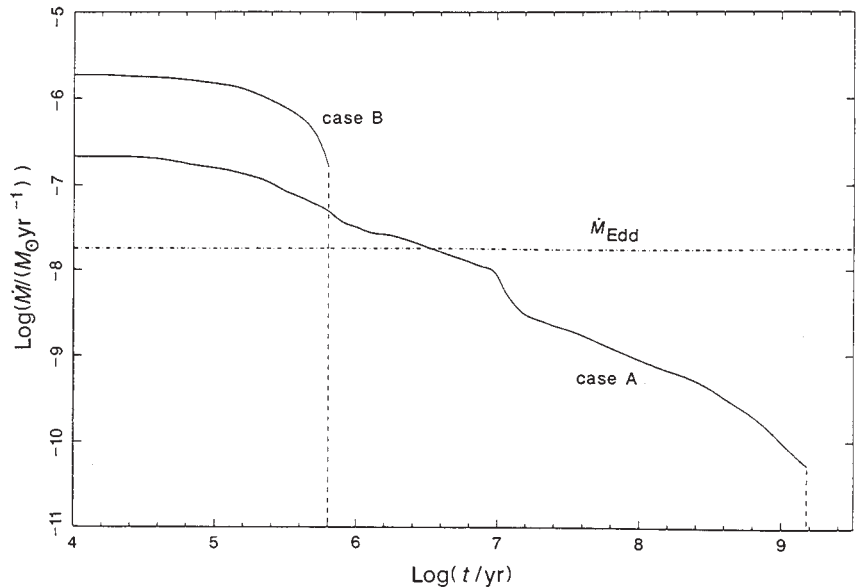
We are now in a position to discuss, at least qualitatively, how irradiation effects modify the evolution of LMXBs. In general, we can distinguish four phases: a turn-on phase, a relaxation phase, a quasi-equilibrium phase and a turn-off phase. The turn-on phase will be modified because of a strong positive feedback between the initial expansion and the external irradiation. The most important new feature, however, is the existence of a relaxation phase, in which mass transfer is driven by the thermal relaxation of the secondary's envelope. The secondary expands during this phase, so the orbital period will generally also increase (provided that the secondary can maintain Roche-lobe contact). After the secondary has achieved its new larger equilibrium configuration, mass transfer would stop if there were no other mechanisms to drive it. If there are other driving mechanisms, such as gravitational radiation and/or magnetic braking, mass transfer may continue, but the further evolution will now be governed by these other mechanisms and will be qualitatively similar to the case where irradiation effects are not

included (in particular, the orbital period may now decrease; see, for example, ref. 7). The main difference is that the secondary and the orbital period will be larger, because the secondary passes through a different sequence of quasi-equilibrium models (see Fig. 1). Unlike the case without external irradiation, however, mass transfer will stop abruptly when the external irradiation flux falls below a certain critical flux $F_{\text{cr}} \sim F_{\text{Ch}}$. This occurs as a result of recombination of hydrogen at the bottom of the irradiated layer, whereupon the secondary starts to shrink more rapidly than the orbit can, because of angular momentum losses from the system, and the secondary becomes detached from its Roche lobe.

Unfortunately, the details of the evolution of the LMXB phase are very sensitive to the evolutionary state of the secondary and to the details of the angular-momentum loss mechanisms, and not all of these phases may be realized for any particular set of assumptions. We have calculated the evolution of two typical systems: in case A, the secondary was an unevolved main-sequence star of mass $M_2 = 0.8 M_\odot$; in case B, it was a subgiant of the same mass. In these calculations, we assumed that the luminosity of the neutron-star primary, which had an initial mass $M_1 = 1.4 M_\odot$, was equal to its accretion luminosity, where the accretion rate was not allowed to exceed the Eddington accretion rate for a neutron star (if the secondary's mass loss rate exceeded the Eddington rate, we assumed that the excess mass was lost from the system). We also included angular-momentum losses resulting from gravitational radiation, magnetic braking (using the heuristic model of Verbunt and Zwaan⁸), and mass loss from the system.

In Fig. 2 we show the mass-loss rate of the secondary as a function of time after the beginning of the mass-transfer phase. In case A (lower solid curve), our results show many similarities to previous calculations that did not include irradiation effects (see for example ref. 7). The main differences are that the orbital period increases during the first $\sim 10^8$ yr from 5.7 to ~ 10 h, and that mass transfer stops after $\sim 10^9$ yr, when the secondary's mass has decreased to $\sim 0.15 M_\odot$, and the period is ~ 8 h. In case B (upper solid curve), the evolution is altered drastically. The secondary's mass-loss rate always exceeds the neutron star's Eddington accretion rate (dot-dashed curve) by more than an order of magnitude, and lasts for only $\sim 10^6$ yr. When the system becomes detached, the orbital period has increased from 45 h initially to a final period of 90 h, and the mass of the secondary has been reduced to $M_2 \approx 0.2 M_\odot$. Indeed, because of the

FIG. 2 Mass loss rate, $\dot{M}$, of the secondary in an irradiated LMXB as a function of time, t , after the beginning of the mass-transfer phase, for two illustrative binary cases. In case A, the secondary is an unevolved star of mass $0.8 M_{\odot}$; in case B, it is a slightly evolved subgiant of the same mass. The initial orbital periods are 5.7 h and 45 h respectively, and the final periods are 8 h and 90 h respectively. In both cases, the secondary is assumed to be a neutron star with an initial mass of $1.4 M_{\odot}$, and the calculations include angular-momentum losses resulting from gravitational radiation, magnetic braking and mass loss from the system. For comparison, the dot-dashed curve shows the Eddington accretion rate for a neutron star.



secondary's large mass-loss rate, it may be necessary to include other effects, which may further alter the evolution of LMXBs of this type (as will be discussed elsewhere).

Although these calculations are only exploratory, they illustrate how irradiation effects may explain a variety of observed features of LMXBs that can not easily be understood within the present theoretical framework. First, if the secondary is a main-sequence star, there is a phase in which the secondary expands and the orbital period increases. Indeed, the rate at which the orbital period increases in our model calculation is of the same order as is observed in Cyg X-3⁹ and X 1822-371 (ref. 10) ($\dot{P} \approx 0.01 \text{ s yr}^{-1}$). Thus, these systems may provide direct observational evidence for the existence of a relaxation phase. Second, the radius of the secondary and the orbital period of systems with main-sequence secondaries are significantly larger than in calculations that do not include irradiation effects. As a result, systems such as AC211 (4U2127+11) may have main-sequence instead of subgiant secondaries¹¹. (We note that our subgiant calculation seems to be inconsistent with the present characteristics of AC211.) Third, in our calculations mass transfer stops abruptly when the external flux drops below a critical value. This introduces a low-luminosity cutoff in the luminosity distribution of LMXBs (at $\sim 10^{36} \text{ erg s}^{-1}$ in our model calculation), which is consistent with the observed luminosity distribution¹². Such a rapid cutoff may also be required to avoid significant spin-down of the neutron-star primary (for a discussion of this point see, for example, ref. 13). Fourth, when the binary becomes detached, the radius of the secondary is larger by a factor of 2–3 than its equilibrium radius without external illumination. This may explain why, in systems such as the binary millisecond pulsar PSR 1557+20, the secondary significantly underfills its Roche lobe¹⁴. Indeed, in our model calculation with a main-sequence secondary, the system at the end of the mass-transfer phase is very similar to PSR1557+20. Fifth, LMXBs with subgiant secondaries (with periods of less than

~ 30 days) are expected to radiate near the neutron star's Eddington limit (which may explain the high luminosity of systems such as Sco X-1), and the duration of the LMXB phase may be much shorter than has been found in earlier calculations¹⁵. If there is a sufficiently large population of LMXBs with subgiant secondaries, this could solve the apparent statistical discrepancy between the number of LMXBs and millisecond pulsars (see, for example, ref. 16).

We expect irradiation effects to be most important for LMXBs, but they may also lead to observational consequences in cataclysmic variables (CVs). As the irradiation fluxes in CVs are much smaller than in LMXBs, and because the energy spectrum is much softer, the effects will generally not be large. Nevertheless, we estimate that the secondaries in CVs may be inflated by up to 10% as a result of external irradiation. The secondary will abruptly become detached from its Roche lobe when the external flux drops below a critical flux. In principle, this could provide an alternative explanation for the period gap in the period distribution of CVs (see, for example, ref. 17).

Finally, we emphasize that the picture outlined here is very preliminary. There are other effects that have not been included but which may change our conclusions quantitatively as well qualitatively. The most important of these are the effects of aspherical illumination, which are likely to lead to internal non-radial circulation; radiation-induced winds¹, which will be particularly important near the end of and after the semi-detached phase; and shielding effects (for example, owing to a flaring disk), which may introduce a negative feedback and may limit the maximum mass-transfer rate.

Despite these caveats, we conclude that, because of the simplicity of the physical process and the large numbers of potential applications, irradiation effects as discussed in this letter are likely to be important in some, if not many, circumstances, and that the theory of the evolution of low-mass X-ray binaries will not be complete without their inclusion. $\square$

Received 20 November 1990; accepted 28 January 1991

1. Ruderman, M., Shaham, J., Tavani, M. & Eichler, D. *Astrophys. J.* **343**, 292–312 (1989).
2. Arons, J. *Astrophys. J.* **184**, 539–550 (1973).
3. Bath, G. T. *Mon. Not. R. astr. Soc.* **171**, 311–328 (1975).
4. Osaki, Y. *Astr. Astrophys.* **144**, 369–380 (1985).
5. Hameury, J. M., King, A. R. & Lasota, J. P. *Astr. Astrophys.* **162**, 71–79 (1986).
6. Tout, C. A., Eggleton, P. P., Fabian, A. C. & Pringle, J. E. *Mon. Not. R. astr. Soc.* **238**, 427–438 (1989).
7. Rappaport, S., Verbunt, F. & Joss, P. C. *Astrophys. J.* **275**, 713–731 (1983).
8. Verbunt, F. & Zwaan, C. *Astr. Astrophys.* **100**, L7–L9 (1981).
9. van der Klis, M. & Bonnet-Bidaud, J. M. *Astr. Astrophys.* **214**, 203–208 (1989).

10. Hellier, C., Mason, K. O., Smale, A. P. & Kilkeny, D. *Mon. Not. R. astr. Soc.* **244**, 39p–43p (1990).
11. Naylor, T., Charles, P. A., Drew, J. E. & Hassall, B. J. *Mon. Not. R. astr. Soc.* **233**, 285–304 (1988).
12. Long, K. S. & Van Speybroeck, L. P. in *Accretion-Driven Stellar X-Ray Sources* (ed. Lewin, W. H. G. & Van den Heuvel, E. P. J.) 117–146 (Cambridge University Press, 1983).
13. Ruderman, M., Shaham, J. & Tavani, M. *Astrophys. J.* **336**, 507–518 (1989).
14. Djorgovski, S., & Evans, C. R. *Astrophys. J.* **335**, L61–L65 (1988).
15. Webbink, R. F., Rappaport, S. & Savonije, G. J. *Astrophys. J.* **270**, 678–693 (1983).
16. Ray, A. & Kluźniak, W. *Nature* **344**, 415–417 (1990).
17. Patterson, J. *Astrophys. J. Suppl.* **54**, 443–493 (1984).

Plasma fireballs formed by microwave interference in air

Y. H. Ohtsuki* & H. Ofuruton†

* Department of Physics, Waseda University, Okubo 3, Shinjuku-ku, Tokyo 169, Japan

† Tokyo Metropolitan College of Aeronautical Engineering, 53-1, Minamisenju 8, Arakawa-ku, Tokyo 116, Japan

DESPITE many experimental and theoretical studies, a widely accepted explanation of ball lightning has yet to be found. It is possible, however, that ball lightning is created by the interaction of different types of energy in the atmosphere¹. Indeed, the production of a fireball by an electric discharge in an atmosphere containing dilute fuel gases such as propane has been reported². We have also managed to produce several types of fireball by setting off electric discharges in an atmosphere containing aerosol with varying concentrations of ethane and/or methane³. Here we report the production of plasma fireballs in a natural atmosphere by microwave interference. The fireballs exhibited certain properties that match eyewitness observations of ball lightning, such as motion against the wind and the ability to pass through a wall intact.

Kapitza proposed an alternative theory of ball lightning in which he predicted where fireballs might form⁴. He suggested that the most favourable places for the production of ball lightning were regions where radio waves, produced by thunderstorms, reached their highest intensity—such places correspond to the potential antinodes obtained in various interference phenomena.

Babat⁵ used a 1–10-MHz continuous-wave (c.w.) microwave, producing a discharge and hence a fireball at normal atmospheric pressure at distances of several centimetres away from the electrodes. Hamilton⁶ and Richie⁷ used pulsating microwaves and produced ball-lightning-like discharges at lower pressures. Silberg⁸ used a 105-MHz c.w. power oscillator in the air with a Lecher wire and produced a standing-wave electric discharge in the region of an antinode in the electric field along the Lecher wires. Powell and Finkelstein⁹ produced plasma fire from the interference of the electromagnetic waves in a glass tube at normal atmospheric pressure, thereby partly confirming Kapitza's model.

But it has proved very difficult to create in the laboratory a plasma fireball that will float and move for a considerable length of time in a natural atmosphere by means of microwave interference.

A schematic diagram of the apparatus used is shown in Fig. 1. We used a 2.45-GHz, 5-kW (maximum power) microwave oscillator (magnetron) operating continuously in the air, and it was usually turned on when a plasma fireball appeared. The microwave radiation was guided through a rectangular waveguide to a cylindrical cavity either made of copper plate or aluminium punched plate. The cavity was 161 mm in diameter and 370 mm in length, and the cross-section of the waveguide was 55 × 109 mm. The waveguide was connected to the cavity in such a way that the short side of the waveguide was parallel to the centre axis of the cavity. One side of the cavity was covered by a piece of immovable aluminium foil and the other side was covered by a movable aluminium punched plate. The air in the cavity was at normal atmospheric pressure. A video camera was set up to record the plasma fireballs as they appeared in the cavity. The lifetime of the plasma fireball (from its appearance to its disappearance) was determined from the video picture.

We observed various kinds of plasma fires at powers ranging from 1 to 5 kW, all, except the last type described below, under the same experimental conditions. The first type of plasma fire observed was stable and practically stationary, but several times was observed to split into two. The fireball lasted for a few

minutes (depending on the continuity of the interference in the cavity). Figure 2a shows the change in the shape and size of the fireballs observed. Although the electric field was probably very strong just under the waveguide, not all of the fireballs appeared in that area. A standing wave should be set up in the cavity by interference between a propagating wave and a reflected wave, and in this case there should be theoretically six antinodes at which the electric field is strongest. But only one plasma fireball was observed to form at a nodal point, and this corresponded to the point of maximum interference because of a small deviation of the cavity's shape from an ideal cylinder.

The second type of plasma fire produced was a very bright homogeneous flash occupying the whole of the semi-cavity. This flash was white in colour, and very brief (~0.2 s), following which a ball of plasma appeared. This fireball subsequently divided into two, which moved up and down, horizontally and around in the cavity (see Fig. 2b). They were probably moving along intense electric field lines in the cavity, their complicated paths therefore the result of the complex electric-field patterns set up in the cavity. These fireballs were orange and lasted for ~2–3 s.

The third type of plasma fire took the form of flames located at the end of the waveguide and separated from the cavity wall (Fig. 2c). The colour of the inside of the cavity changed from white and red to blue and orange. The size of the flame varied from 1 to 3 cm, and it lasted for a few seconds only. This type of plasma fire often moved into the waveguide from the cavity. When a 3-mm-thick ceramic board was inserted between the cavity and the waveguide, the plasma fireball moved through the ceramic board intact without causing any damage or melting. This phenomenon is reminiscent of reports of ball lightning moving through a wall or roof.

The fourth type of plasma fire observed was, like the first type, stable. But it flowed up out of the cavity into free space and the whole cavity was illuminated by the discharge. The plasma fireball burnt a hole in the aluminium foil that covered one side of the cavity. A few seconds later, the plasma fire emerged from a hole in the aluminium foil (see Fig. 3). This plasma fire lasted for ~1–2 s, even after the magnetron was turned off, but its occurrence was very rare (usually it was necessary to continue supplying energy to produce fireballs that lasted a considerable length of time). This phenomenon may be the result of the large time constant in the decay of energy stored in the cavity; however, we can find no reasonable explanation for the plasma fire lasting for such a long time after the magnetron has been turned off. The visible plasma flame reaching through the hole was not caused by combustion of the aluminium foil, because the hole did not become any bigger while the plasma flame was visible.

In our final experiment, a copper bar (100 mm in length) was placed just under the waveguide in the cavity. We then observed

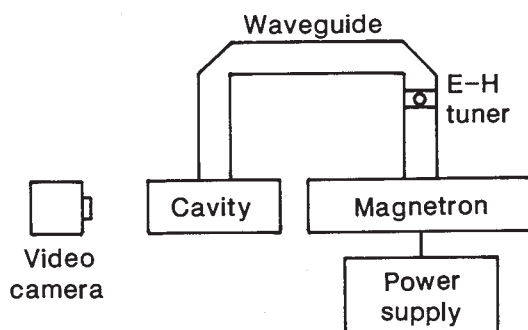


FIG. 1 Schematic diagram of the apparatus. The cavity is 161 mm in diameter and 370 mm in length. The cross-section of the waveguide has dimensions 55 × 109 mm.

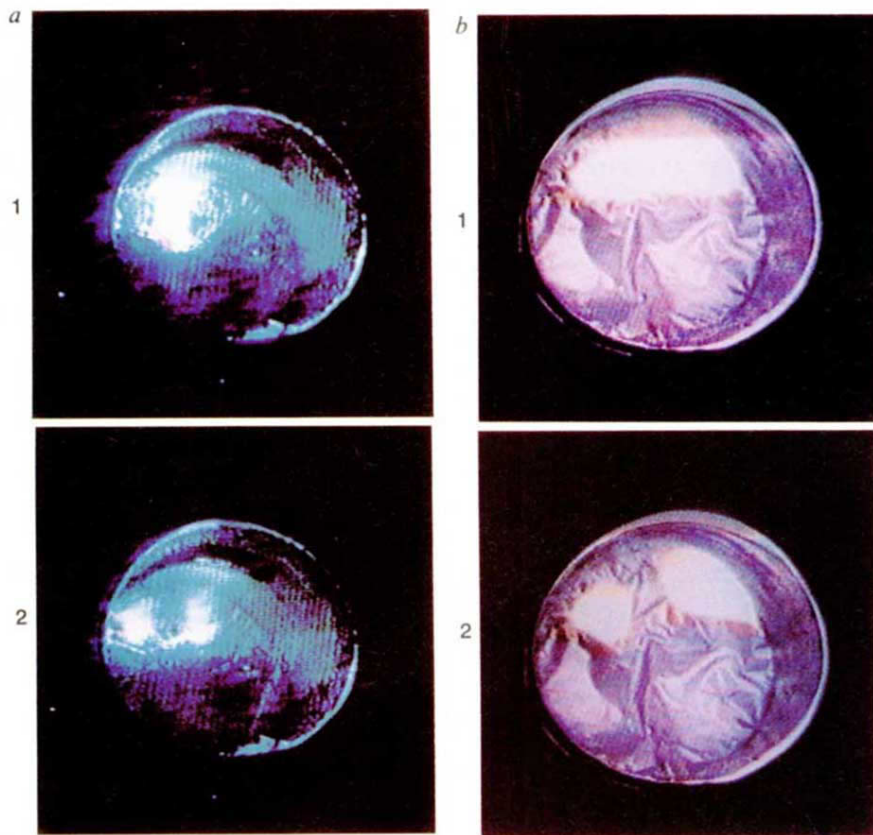


FIG. 2 *a*, First type of plasma fire. This type sometimes split into two as shown in frame 2. Frame 1, 1 s after the plasma fire appeared; Frame 2, 5 s after the plasma fire appeared. *b*, Second type of plasma fire. This plasma fire moved up and down in the cavity. Frame 1, 0.12 s after the plasma fire appeared; Frame 2, 0.33 s after the plasma fire appeared. *c*, Third type of plasma fire. The colour in the cavity changed from white and red to orange and blue. In these experiments the cavity was made of copper plate.

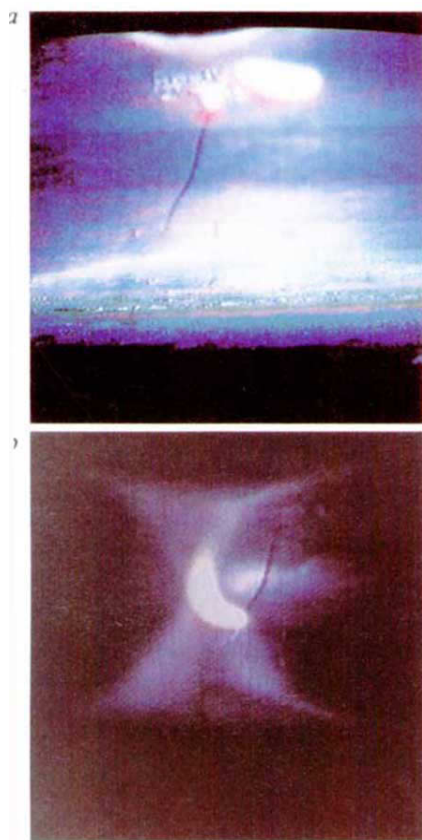


FIG. 4 Fifth type of plasma fire: *a*, 5 s after the plasma fire appeared; *b*, 20 s after the plasma fire appeared. The cavity was made of aluminium punched plate.

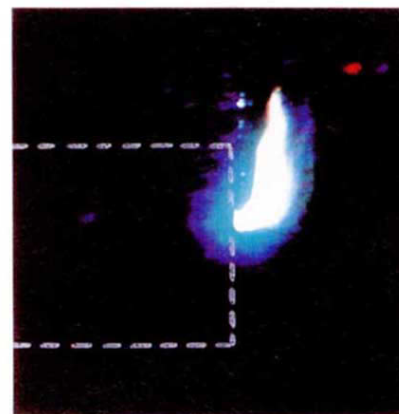


FIG. 3 Fourth type of plasma fire. This type blew out of the cavity. White dashed lines indicate the outline of the cavity (made of aluminium punched plate).

this fifth type of plasma fire at the top of the bar. Many plasma balls appeared one after another, but the lifetime of each was very short (~ 1 s). Some of them travelled down along the copper bar, and some of them became detached. Interestingly, many witnesses of ball lightning have reported seeing it travel along an electric wire (see Fig. 4). Moreover, when a strong wind from a compressor was blown into the cavity, some plasma balls moved against the wind, indicating that the fireballs will move along an intense electric field even if there is an opposing flow of air. None of the plasma fireballs formed in our experiments was in contact with the cavity wall.

Electromagnetic wave interference is rarely observed in the natural world, and our experiments do not therefore elucidate

the process of excitation responsible for natural ball lightning or the energy source for its production. There have, however, been a few studies of the effect of electromagnetic radiation in a natural atmosphere^{12,13}, and recently Zheng¹⁴ made a calculation of the radiation intensity required to produce a fireball effect. But further theoretical and experimental studies are needed. We believe that electromagnetic interference might be found in the natural atmosphere where the topography is complex and where there is rapid variation in atmospheric electricity. The analogies that we found between the properties of experimental fireballs and those reported for ball lightning seem to confirm the role of such electromagnetic interference.

We are now attempting to use various kinds of semi-cavity, with a varying composition of aerosol, vapour and fuel gases. We also intend to examine Handel's 'caviton model'¹⁵ for ball lightning in our experiments. □

Received 27 July 1990; accepted 7 January 1991.

1. *Science of Ball Lightning (Fire Ball)* (ed. Ohtsuki, Y. H.) (World Scientific, Singapore, 1989).
2. Barry, J. D. *J. atmos. terr. Phys.* **30**, 313–317 (1968).
3. Ofuruton, H. & Ohtsuki, Y. H. *Nuovo Cim.* **130**, 761–768 (1990).
4. Kapitza, P. L. *Dokl. Akad. Nauk. SSSR* **101**, 245–248 (1955).
5. Babat, G. I. *J. Inst. electr. Eng.* **94**, Part 3, 27–37 (1947).
6. Hamilton, C. W. *Nature* **188**, 1098–1099 (1960).
7. Richie, D. J. *J. Inst. electr. Eng.* **9**, 202–206 (1963).
8. Silberg, P. A. *J. appl. Phys.* **35**, 2264–2265 (1964).
9. Powel, J. R. & Finkelstein, D. *Am. Sci.* **58**, 262–280 (1970).
10. Brand, W. *Der Kugelblitz* (Henri Grand, Hamburg, 1923).
11. Barry, J. D. *Ball Lightning and Bead Lightning* (Plenum, New York, 1980).
12. Kosarev, E. L. et al. *Soviet Phys. Tech. Phys.* **13**, 1477–1479 (1969).
13. Kosarev, E. L. & Serezhkin, Yu. G. *Soviet Phys. Tech. Phys.* **19**, 229–233 (1974).
14. Zheng, X. H. *Phys. Lett. A* **148**, 463–469 (1990).
15. Handel, P. H. in *Science of Ball Lightning (Fire Ball)* (ed. Ohtsuki, Y. H.) 254–259 (World Scientific, Singapore, 1989).

ACKNOWLEDGEMENTS. The authors thank Prof. K. Kohra and Mr. H. Ara for their help. We also thank Tsukuba Institute of Science and Technology.

The 154-day and related periodicities of solar activity as subharmonics of a fundamental period

T. Bai & P. A. Sturrock

Center for Space Science and Astrophysics, Stanford University, Stanford, California 94305, USA

IN 1984 a periodicity of 154 days was found in the record of solar flare activity¹ from 1980 to 1983. Since then, the same periodicity has been found in many different measures of solar flare activity during cycle 21 (soft X-ray peak flux¹; hard X-ray emission^{2,3}; H α flare activity⁴; microwave peak flux⁵; production of interplanetary electrons⁶ and protons⁷; 10-cm radio flux⁸) as well as in sunspot area^{9,10}. The cause of this 154-day periodicity, which has also been discovered in solar cycles 19 and 20 (refs 7, 11) remains unknown, but a suggestion that it was related to enhanced flare activity in certain longitude bands has been ruled out³. Here we show that periodicities of 51, 78, 104 and 129 days, in addition to the 154-day period, can often be detected in flare and sunspot records. These periods are close to integral multiples (by factors of 2, 3, 4, 5 and 6) of 25.8 days, suggesting that they are subharmonics of a fundamental period.

Although there is evidence that the 154-day periodicity was in operation during solar cycle 20, in most cases, for the same kind data, the 154-day power for cycle 20 is weaker than the 154-day power for cycle 21 (refs 4, 5, 8, 11). The strongest evidence that the 154-day periodicity was in operation during cycles 19 and 20 is found in the rate of flares that produce

interplanetary energetic protons⁷. When the proton-flare rates of cycles 19 (1955–1965) and 20 (1966–1975) are analysed together, it is found that the 154-day power comes mainly from the 1958–1971 interval. The 154-day peak for this interval is at a 99.8% confidence level. Bai and Cliver⁷ showed that the phase of the 154-day periodicity had shifted by about 180° when the periodicity resumed in the proton-flare occurrence rate for the interval from April 1978 to August 1983. (See ref. 12 for a discussion of phase coherence.)

There is evidence from analyses of the daily sunspot number¹² and the daily sunspot area (ref. 13; and M. Carbonell and J. L. Ballester, preprint) that the 154-day periodicity was present before solar cycle 19. In addition, the power spectrum of the equatorial solar diameter measured during the interval 1683–1718 shows 6 σ peak at 155 days¹⁴. Thus it seems that the appearance of this periodicity is not a one-time chance event but represents a fundamental property of the Sun. The values of period found by various authors fall in the range from 150 to 158 days.

We calculated the statistical significance of periodicities as follows. The distribution function of power follows an exponential form at low values of power, and when the power is properly normalized, the probability p of obtaining a peak with a height h at a given frequency by chance is simply $p = \exp(-h)$ (ref. 7 and references therein). The probability of obtaining such a high peak at one or more of N independent frequencies is

$$p = N \exp(-h) \quad (1)$$

Here $N = (f_2 - f_1)T$, where $(f_2 - f_1)$ is the frequency interval of search and T is the time interval for the data run.

To determine the time interval during which the periodicity operates, we use the maximum-likelihood method. The logarithmic likelihood function is defined as

$$m(A, P, \phi; n) = \sum_{i=1}^n X_i \ln [1 + A \cos 2\pi(t_i/P + \phi)] \quad (2)$$

where X_i is the time series normalized such that the average is unity, P is period and ϕ is phase. Using the power spectral analysis, we can determine the period and the phase angle, from which in turn we can determine the amplitude A that maximizes the likelihood function. Once we know the values of A , P and ϕ that maximize the likelihood, we can study the logarithmic likelihood as a function of n (time), because it is cumulative. In a cycle of period P for which the signal displays the periodic modulation, the X_i s are likely to be greater than 1 when $\ln [1 + A \cos 2\pi(t_i/P + \phi)] > 0$ (or equivalently,

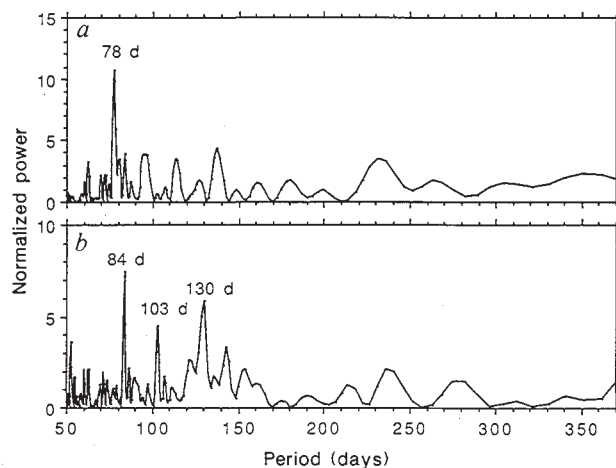


FIG. 1 *a*, Normalized power spectrum of daily sunspot-area measurements from 10 November 1971 to 14 July 1978. *b*, Normalized power spectrum of the rate of flares with CFI > 6 for solar cycle 20.

$[1 + A \cos 2\pi(t_i/P\phi)] > 1$, and X_i are likely to be smaller than 1 when $\ln[1 + A \cos 2\pi(t_i/P + \phi)] < 0$. Therefore, when a cycle exhibiting the periodicity is completed, the logarithmic likelihood function increases; otherwise it decreases. We identify an interval of consecutive cycles for which the logarithmic likelihood function increases as an interval for the periodicity. This is an objective way of finding time intervals for which the periodicity is in operation. This method is a generalization of the method given by Bai and Cliver⁷.

Bai¹⁵ analysed the rate of flares selected by the comprehensive flare index (CFI) from compilations of Dodson and Hedeman^{16,17} for cycle 19, and found that the strongest periodicity was at 50.9 (± 0.7) days. (Hereafter the range of period indicates the full width at half maximum.) This periodicity is significant at a 98.5% level¹⁵. We find that the 51-day periodicity was evident from November 1956 to September 1960. The 51-day periodicity is found to be as strong as the 154-day periodicity in the power spectrum of daily sunspot-area measurements for cycle 21 (ref. 18). Also there is a 51-day peak in the power spectrum of the whole-disk 10-cm radio flux measurement for the interval 1948–1987⁹. Furthermore, in the power spectrum of the γ -ray flare rate for cycle 21 (ref. 6), the 51-day peak is the strongest after the 154-day peak in the period range 23–1,000 days. As 51 days is about one-third of 154 days, a connection between the two periodicities has been proposed¹⁵.

In the power spectrum of the rate of flares selected by microwave peak fluxes for solar cycle 20, the peak at 78 days is the strongest in the 30–500-day interval⁵. During the 1971–1977 interval, the proton flare rate did not exhibit the 154-day periodicity (compare ref. 7). We analysed the power spectrum of the daily sunspot-area measurements and found a strong peak at 77.8 (± 1.0) days (Fig. 1a), which is half of 154 days. The 78-day peak has a power of 10.7. The probability of obtaining such a peak by chance at one of the 200 frequencies that we studied is $P = 200 \exp(-10.7) = 0.005$. The 78-day periodicity is found to be present from November 1971 to June 1978, roughly the same interval as that during which the proton flare rate did not exhibit the 154-day periodicity. Thus, in late 1971 the 154-day periodicity switched to a periodicity with half the period, and then in 1978 it switched back to the 154-day periodicity. The time interval during which the 154-day periodicity was switched to the 78-day periodicity corresponds to 31 cycles of 78 days. This odd number of cycles may be the reason for the 180° phase shift of the 154-day periodicity of the proton flare rate.

Figure 1b shows the normalized power spectrum of the rate of flares with CFI > 6 for cycle 20 (flare compilations of refs 17, 19). The three strongest peaks are at 84, 130 and 103 days, in that order. The power for the 84-day periodicity is found to come mainly from the interval from August 1965 to December

1975. The peak is at 83.8 days. The power for the 130-day periodicity comes mostly from the two intervals November 1967 to January 1971 (9 cycles) and March 1973 to July 1978 (15 cycles), and if we confine our analysis to the interval from November 1967 to July 1978, the peak is at 129.9 (± 0.7) days. The power for the 103-day periodicity comes mainly from the interval between November 1969 and February 1976, and if we confine our analysis to this interval, the peak is at 104.4 (± 1.6) days. The normalized power of the peak in Fig. 1b is 4.55; the probability of obtaining such a peak by chance at a given frequency is $p = \exp(-4.55) = 0.01$. The probability of obtaining such a peak anywhere between 30 and 230 nHz (50–370 days) is large, but the probability for finding such a peak at a value near four times 25.8 days (in the range 100–105 days, two-thirds of the range 150 to 158 days) is only 6%. The 130-day peak has a normalized power of 5.91, and the probability of obtaining such a peak near five times 25.8 days (125–132 days) is about 1%.

Figure 2 summarizes the power spectral analysis of CFI flares (1955–1979), SMM flares (flares with peak rates > 1,000 counts per s, measured by hard X-ray on the Solar Maximum Mission, 1980–1984), daily sunspot-area measurements (1971–1985) and proton flares (1955–1985). These data sets, which cover all or parts of cycles 19–21, have been analysed in detail recently^{3,7,9,16}. Proton flares cover all three cycles, and CFI flares and SMM flares together cover three solar cycles. We cannot help but notice that, except for 84 days, all the periods are equal to integral multiples of 25.8 days within the error bars quoted earlier. Integral multiples of the 25.0–26.3-day interval cover about 5% of the 50–370-day interval. The probability that four out of the five additional periodicities (in addition to the 154-day period) in Fig. 2 fall by chance in these narrow intervals is $p = 5 \times (0.05)^4 = 3.1 \times 10^{-5}$.

A nonlinear dynamical system with a periodic forcing term can show periodic behaviour at subharmonic frequencies (that is, at periods that are integral multiples of the fundamental periods). For instance, one of the best-studied nonlinear systems is that described by Duffing's equation,

$$\frac{d^2x}{dt^2} + k \frac{dx}{dt} + x^3 = B \cos \omega t \quad (3)$$

where k and B are constants. For certain values of k and B , this system can exhibit periodic behaviour at subharmonics, and for other values it can show chaotic behaviour^{20,21}.

According to this interpretation, solar activity involves a nonlinear response to a solar 'clock' with a period of about 25.8 days. Determining the nature of this clock then becomes the key to understanding the 154-day and related periodicities exhibited by solar activity. Evidence for the 25.8-day clock and its interpretation will be discussed elsewhere²². □

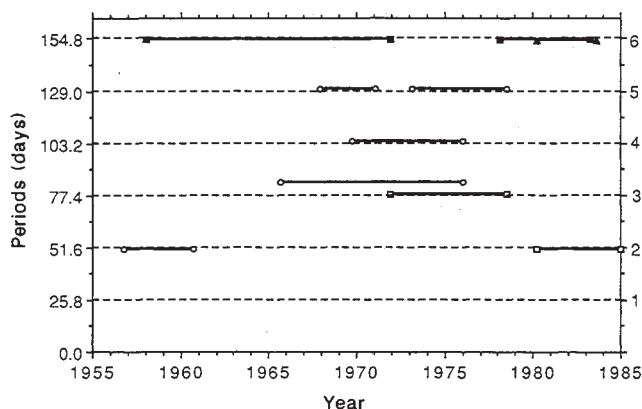


FIG. 2 Various periodicities as integral multiples of 25.8 days. The horizontal bars indicate the time intervals for various periods. ■, Proton flares; ▲, SMM flares; ○, CFI flares; □, sunspot area.

Received 23 October 1990; accepted 17 January 1991.

1. Rieger, E. *et al. Nature* **312**, 623–625 (1984).
2. Dennis, B. R. *Solar Phys.* **100**, 465–490 (1985).
3. Bai, T. & Sturrock, P. A. *Nature* **327**, 601–604 (1987).
4. Ichimoto, K., Kubota, J., Suzuki, M., Tohmura, I. & Kurokawa, H. *Nature* **316**, 422–424 (1985).
5. Bogart, R. S. & Bai, T. *Astrophys. J.* **299**, L51–L55 (1985).
6. Droege, W., Gibbs, K., Grunfeld, J. M., Meyer, P. & Newport, B. J. *Astrophys. J. Suppl.* **73**, 279–283 (1990).
7. Bai, T. & Cliver, E. W. *Astrophys. J.* **363**, 299–309 (1990).
8. Kile, J. N., Cliver, E. W. & Fougere, P. F. *Astrophys. J.* (in the press).
9. Pap, J., Tobsika, W. K. & Bouwer, S. D. *Solar Phys.* **129**, 165–189 (1990).
10. Lean, J. L. & Brueckner, G. E. *Astrophys. J.* **337**, 568–578 (1990).
11. Ozguc, A. & Atac, T. *Solar Phys.* **123**, 357–365 (1989).
12. Wolff, C. L. *Astrophys. J.* **264**, 667–676 (1983).
13. Lean, J. L. *Astrophys. J.* **363**, 718–727 (1990).
14. Ribes, E., Meriin, Ph., Ribes, J.-C. & Bartholot, R. *Annales Geophys.* **89-04**, 321–329 (1989).
15. Bai, T. *Astrophys. J.* **318**, L85–L91 (1987).
16. Dodson, H. W. & Hedeman, E. R. *World Data Center for Solar-Terrestrial Physics Report UAG-14* (1971 NOAA, Boulder).
17. Dodson, H. W. & Hedeman, E. R. *World Data Center for Solar-Terrestrial Physics Report UAG-52* (NOAA, Boulder, 1975).
18. Bouwer, S. D., Donnelly, R. F. & Pap, J. in *Proc. XXVIII COSPAR* (Hague, The Netherlands, in the press).
19. Dodson, H. W. & Hedeman, E. R. *World Data Center for Solar-Terrestrial Physics Report UAG-80* (NOAA, Boulder, 1981).

20. Thompson, J. M. T. & Stewart, H. B. *Nonlinear Dynamics and Chaos* (Wiley, New York, 1986).
 21. Ueda, Y. in *New Approaches to Nonlinear Problems in Dynamics* (ed. Holmes, J. P.) 311–322 (Philadelphia, SIAM, 1981).
 22. Sturrock, P. A. & Bai, T. *Bull. Am. astr. Soc.* **22**, 1225 (1990).

ACKNOWLEDGEMENTS. This research was supported by NASA and the Office of Naval Research.

Latex elastomer with a permanently hydrophilic surface

Isao Noda

The Procter & Gamble Company, Miami Valley Laboratories,
 PO Box 398707, Cincinnati, Ohio 45239-8707, USA

WATER wettability is important in many applications of polymeric materials, including fabrics, printing and biomedical uses^{1,2}. Various surface-modification techniques are available to convert the intrinsically hydrophobic surfaces of plastic to water-wettable ones, by incorporating chemically polar functional groups at the surface^{4–7}. Although such chemically induced surface hydrophilicity can be relatively long-lived while the substrates remain rigid, increases in the mobility of surface molecules, for example, due to increasing temperature, can cause rapid loss of hydrophilicity, driven by the tendency of surfaces to minimize their free energy^{2–5}. As the polar groups are not often bound to the polymer matrix (usually being instead free surfactant molecules), they may also be flushed from the surface by repeated exposure to water. Here I report the preparation of permanently water-wettable elastomeric films from a latex synthesized by polymerization of monomers in the presence of an amphiphilic block copolymer. By migration of the hydrophilic segments to the surface during film formation, the film is rendered essentially completely wettable by water. Applications may include flexible coatings that can introduce spatially selective wettability to solid surfaces, for example one-side wettable perforated films used for bandages and disposable diapers⁸.

Figure 1a shows a water droplet placed on a film of synthetic rubber. The water contact angle exceeds 90°, indicating the film surface is hydrophobic. Similar water-repelling behaviour is observed for most natural or synthetic rubbers, unless surfactants are added to the system to lower the surface tension of the water. Such reduction of the liquid surface tension certainly results in an improvement of the apparent wettability of the surface, but surfactant-induced wetting is in many ways qualitatively different from the spreading of water over truly hydrophilic solid surfaces.

In particular, the effect of surfactant treatment is usually temporary. The wettability diminishes once the surface is rinsed with excess water. Migration of surfactants away from the surface to the bulk interior or to other nearby surfaces can also occur. Furthermore, in many applications (such as printing, coating and biomedical applications), it is often desired to make only a part of the available solid surfaces water-wettable while keeping the rest of the system dry. For such selective, spatially resolved wetting, the distinction between facilitating wetting by making the surface intrinsically hydrophilic or by simply reducing the liquid's surface tension becomes very important.

Figure 1b shows a water droplet on the permanently hydrophilic surface of a film made from the new latex that I have developed. The latex was prepared by polymerizing a mixture of monomers (40 parts styrene to 60 parts butadiene) dispersed in water. An amphiphilic block copolymer, consisting of a polyethylene oxide segment of relative molecular mass (M_r) 1,365 attached to a short-chain polybutadiene (M_r = 480), was used as an emulsifying agent instead of a typical low-molecular-weight surfactant⁹. Each latex particle is believed to be covered by a layer of the block copolymer, and a substantial fraction of the reactive polybutadiene segment is probably grafted to the

latex core during the emulsion polymerization. The latex exhibits an exceptionally high colloidal stability. Extensive dialysis of the latex, even for days, diminishes neither the colloidal stability nor the capability of the latex to form a hydrophilic film.

When the latex is dried, the soft particles fuse together to form a continuous rubbery film. During this stage, redistribution of the block copolymer apparently occurs such that there is a substantial surface excess concentration of the ethoxylate groups (this was detected by an infrared technique). There is a corresponding marked increase in the surface hydrophilicity of the resulting rubber film. A water droplet placed on the film readily spreads over the entire surface (Fig. 1b). The advancing water contact angle is so small (well below 10°) that a precise measurement is difficult. Practically speaking, the rubber film is completely water-wettable.

Although the driving force for the surface accumulation of the amphiphilic block copolymer has not yet been identified, it is likely to be of entropic origin associated with the large size of the polymeric amphiphile, which somehow compensates for the energetic penalty of bringing the polar, high-energy component to the surface. Ageing of the surface-hydrophilic rubber film in dry air at room temperature for several months did not alter the water wettability. The surface hydrophilicity is equally stable at elevated temperatures (>60 °C). If the surface hydrophilicity is due to a simple kinetic effect associated with trapping of hydrophilic constituents near the surface during the formation of the film, prolonged exposure to dry air should eventually drive such moieties away from the surface into the bulk phase^{2–5}. The surface hydrophilicity of this material, however, seems quite stable.

Several experiments were conducted to show that the observed water wettability indeed arises from the increased surface hydrophilicity of the rubber film and not from the lowering of the surface tension of the water droplet by residual surfactants. Rinsing with excess running water is usually sufficient to remove most of the surfactant from the surface of an ordinary rubber,

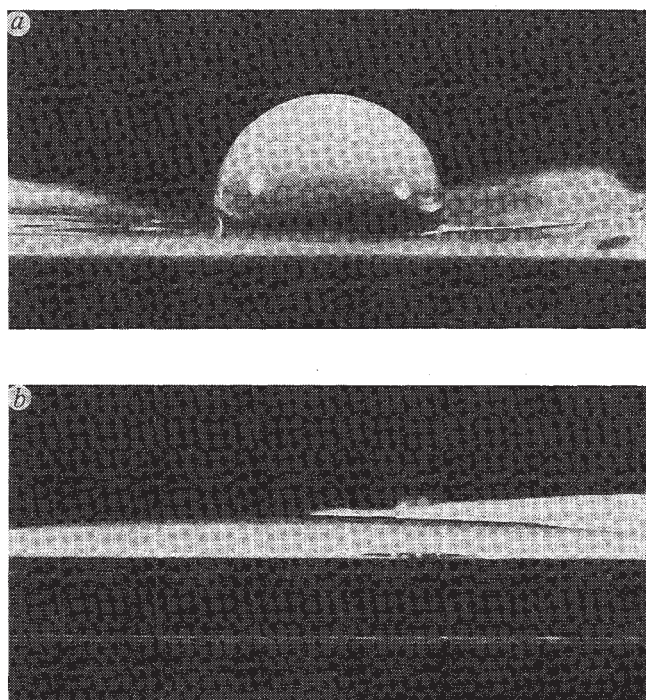


FIG. 1 a, A droplet of water placed on a film of ordinary styrene-butadiene rubber. The water beads up, indicating that the surface is hydrophobic. b, Water droplet spreading over the hydrophilic surface of the new elastomer film.

making it hydrophobic. This surface-hydrophilic rubber film, however, remained completely water-wettable even after many days of rinsing. The amphiphilic block copolymer responsible for the wettability is attached quite firmly (probably covalently) to the rubber surface. The bulk of the material remained totally hydrophobic: there was no detectable swelling of the film by water.

Further evidence that the amphiphilic block copolymer is attached to the rubber surface and not dissolved in the water is obtained by transferring water that has been spread over the hydrophilic rubber onto an ordinary hydrophobic film. The water beads up in a manner similar to that shown in Fig. 1a, indicating that the surface tension of the water has not been substantially reduced. When it is placed back on the hydrophilic rubber, the same water spreads again. This procedure can be repeated many times, and demonstrates that the spreading of water over the rubber film is due to the surface hydrophilicity of the film rather than lowering of the liquid's surface tension. □

Received 3 December 1990; accepted 17 January 1991.

1. Ruckenstein, E. & Gourisankar, S. V. *J. Colloid Interface Sci.* **101**, 436–451 (1984).
2. Ratner, B. D., Weathersby, P. K., Hoffman, A. S., Kelly, M. A. & Scharpen, L. H. *J. appl. Polymer. Sci.* **22**, 643–664 (1978).
3. Sharma, A. K., Millich, F. & Hellmuth, E. W. *J. appl. Polymer. Sci.* **26**, 2205–2210 (1981).
4. Yasuda, H., Sharma, A. K. & Yasuda, T. *J. Polymer Sci. Polymer Phys. Ed.* **19**, 1285–1291 (1981).
5. Briggs, D., Rance, D. G., Kendall, C. R. & Blythe, A. R. *Polymer* **21**, 895–900 (1980).
6. Nuzzo, R. G. & Smolinski, G. *Macromolecules* **17**, 1013–1019 (1984).
7. Rasmussen, J. R., Stedronsky, E. R. & Whitesides, G. M. *J. Am. chem. Soc.* **99**, 4736–4745 (1977).
8. Noda, I. US Patent 4,735,843 (1988).

Ice-core record of oceanic emissions of dimethylsulphide during the last climate cycle

M. Legrand*, C. Feniet-Saigne*, E. S. Saltzman†, C. Germain†, N. I. Barkov‡ & V. N. Petrov‡

* Laboratoire de Glaciologie et Géophysique de l'Environnement, BP 96, 38402 St Martin d'Hères, Cedex, France

† Rosenstiel School of Marine and Atmospheric Science, University of Miami, 4600 Rickenbacker Causeway, Miami, Florida 33149, USA

‡ Arctic and Antarctic Research Institute, 38 Beringa Street, Leningrad 199226, USSR

THE Vostok ice core in Antarctica has provided one of the longest climate records, enabling the stable-isotope, major-ion and gas composition of the atmosphere to be reconstructed over many thousands of years. Here we present depth profiles along this core of methanesulphonate and non-seasalt sulphate (produced by the atmospheric oxidation of dimethylsulphide), which provide the first historical record of biogenic sulphur emissions from the Southern Hemisphere oceans over a complete glacial-interglacial cycle (160 kyr). Those measurements confirm and extend some previous observations made on a very limited data set from the Dome C ice core in Antarctica, which indicated increased oceanic emissions of dimethylsulphide during the later stages of the glacial period, compared with the present day¹. The observed glacial-interglacial variations in methanesulphonate and non-seasalt sulphate confirm that the ocean-atmosphere sulphur cycle is extremely sensitive to climate change.

During initial attempts to remove surface contamination from the outer part of the ice-core sections, we washed the core with ultrapure water at a temperature of ~20 °C. This method has been used successfully for the analysis of mineral species², and to study methanesulphonate (MSA) content over the past 30 kyr in the Dome C ice core¹. Vostok data obtained using this

TABLE 1 Non-seasalt sulphate and MSA in the Vostok and Dome C ice cores

	Dome C		Vostok	
	MSA (ng g ⁻¹)	n.s.s. SO ₄ ²⁻ (ng g ⁻¹)	MSA (ng g ⁻¹)	n.s.s. SO ₄ ²⁻ (ng g ⁻¹)
Holocene (Stage A, 0–10 kyr BP)	2.5 (W)	70 (W)	4 (W) 5 (L)	108 (W) 102 (L)
Glacial age (Stage B, 18–30 kyr BP)	13 (W)	136 (W)	5 (W) 31 (L)	230 (W) 222 (L)

W, data obtained by washing the ice-core sections. L, data obtained by subcoring with a lathe (see text). Vostok data, this work; Dome C data, ref. 2.

procedure are compared with Dome C data for the same climate periods (Table 1). Although the non-seasalt (n.s.s.) sulphate increases from stage A to stage B by a factor of two in both cases, the MSA content at Vostok remains fairly constant during the climate change whereas the Dome C core exhibits an increase by a factor of five from the warm to the cold stage. Such a difference is difficult to understand because Vostok (located 600 km from Dome C on the high Antarctic plateau) has similar geographical and meteorological characteristics (elevation, mean annual temperature). Laboratory experiments indicate that extensive rinsing of Vostok ice leads to mobility of MSA, which does not occur with other ions. This effect may be related to the fact that the Vostok core was drilled with kerosene drilling fluid, whereas the Dome C core was drilled dry. In addition, the Vostok core sections were stored for 24 hours at –3 °C before washing, to prevent breakage of the core.

Subsequently, we designed an ice-core lathe equipped with a stainless-steel knife to mechanically remove the outer part of the core (until 95% of the volume was removed). The MSA content of successive fractions from the outer to the inner part of the core (not reported here) is constant, indicating the reliability of the procedure. Using this method, the increase in MSA during the glacial age in the Vostok core (Table 1) is similar in magnitude to that at Dome C. We prepared 47 ice sections in this manner, each representing approximately six years of accumulation. These sections were cut from the core at roughly 40-m intervals over the length of the 2,080-m core³.

For each sample, the concentrations of Na⁺, Ca²⁺, SO₄²⁻ and CH₃SO₃⁻ were determined by ion chromatography. MSA and sulphate were determined using an AS5 separator column, a gradient pump system (Dionex 4000i) and an integrator with an accuracy of 5% (at the 95% confidence level). Cations were measured using working conditions reported elsewhere⁴. In a previous Vostok ice-core study⁵, it was established that the calcium content mainly reflects the terrestrial input to the Vostok ice and that it is feasible to estimate the marine component of Na⁺ (denoted Na_m⁺) from determinations of Ca²⁺ and Na⁺. It is then possible to calculate the content of non-seasalt sulphate by subtracting seasalt sulphate from the total sulphate.

The profile of n.s.s. sulphate (Fig. 1b) shows moderate glacial-interglacial changes from 98–102 ng g⁻¹ during interglacial stages (G and A, respectively) to 208–240 ng g⁻¹ during cold stages (B, D and H) (see Table 2). This picture is very similar to that from previously published data^{5,6}. Between 1,200 and 1,300 m three samples exhibited unusually high levels of n.s.s. sulphate (>400 ng ml⁻¹). By comparison with data previously obtained at similar depth⁵, it seems that these high concentrations are not related to the long-term pattern of sulphate deposition. Instead, they reflect short-term changes associated with volcanic eruptions which sporadically perturb the background level of sulphuric acid in the Vostok precipitation⁷. At these three levels no significant changes in the MSA content is observed. We therefore neglect these values here. The variations in the MSA profile, like those of n.s.s. sulphate, strongly correlate with climate conditions, changing from 5 ng g⁻¹ during inter-

TABLE 2 Mean concentrations of MSA and n.s.s. SO_4^{2-} in the Vostok core

Stage	Depth (m)	MSA (ng g^{-1})	n.s.s. SO_4^{2-} (ng g^{-1})	R (%)	MSA' (ng g^{-1})	n.s.s. SO_4^{2-} ' (ng g^{-1})	S'	Q _s
A	0-275	5 $\pm$ 2	102 $\pm$ 11	5 $\pm$ 1.5	5 $\pm$ 2	102 $\pm$ 11	107 $\pm$ 13	1
B	325-550	31 $\pm$ 3	222 $\pm$ 7	14 $\pm$ 1.5	20 $\pm$ 2	146 $\pm$ 5	166 $\pm$ 5	1.55
C	575-850	22 $\pm$ 6	168 $\pm$ 14	13 $\pm$ 4	16 $\pm$ 4	120 $\pm$ 9	136 $\pm$ 10	1.27
D	875-975	32 $\pm$ 2	242 $\pm$ 17	13 $\pm$ 1	21 $\pm$ 1	162 $\pm$ 9	183 $\pm$ 9	1.70
E	1,000-1,475	13.0 $\pm$ 4	139 $\pm$ 20	10 $\pm$ 4	11 $\pm$ 3	109 $\pm$ 18	120 $\pm$ 17	1.1
F	1,500-1,600	19 $\pm$ 3	148 $\pm$ 7	13 $\pm$ 2	13 $\pm$ 2	105 $\pm$ 5	118 $\pm$ 7	1.1
G	1,625-1,875	5.5 $\pm$ 3	98 $\pm$ 17	5.4 $\pm$ 2	5.5 $\pm$ 3	100 $\pm$ 11	106 $\pm$ 13	1.0
H	1,925-2,080	25 $\pm$ 1	208 $\pm$ 50	12.5 $\pm$ 3	16 $\pm$ 1	136 $\pm$ 31	152 $\pm$ 30	1.4

R is the weight ratio of MSA to n.s.s. SO_4^{2-} . Primed values MSA' and n.s.s. SO_4^{2-} ' are concentrations corrected for variations in accumulation rate. S' is the sum of MSA' and n.s.s. SO_4^{2-} ' and Q' is the ratio of S' in each stage to S' in stage A.

glacial stages (A and G) to 25–31 ng g^{-1} during very cold stages (B, D and H).

One of the notable features of the record is the similarity of the general trends in MSA and n.s.s. sulphate. Because oceanic sulphur is an important source of MSA, this covariance strongly suggests that n.s.s. sulphate in the precipitation over the central Antarctic is primarily derived from oceanic emissions. It is also possible that sources of non-marine sulphate covary with the marine component. Given recent evidence regarding productivity limitation by aerosol-transported nutrients such as iron^{8,9},

it is perhaps possible that the marine emissions of dimethylsulphide (DMS) from the southern ocean and the long-range transport of continental-derived sulphate might correlate over long timescales and large variations in climate.

The relative magnitude of the glacial-interglacial changes in MSA are greater than those of n.s.s. sulphate, resulting in a decrease in the ratio R of MSA to n.s.s. sulphate during warming. During cold stages R is high, with values reaching 14%, and during warm stages R reaches a minimum of 3%. In the present atmosphere, there is a systematic trend of R in marine aerosols, from low ratios in the low and mid-latitudes (2–10%, ref. 10) towards higher ratios at higher latitudes (15–100%; ref. 11). This effect is presumably atmospheric, related to the different temperature-dependence of OH addition and H-atom abstraction in the oxidation of DMS by OH (ref. 12). It is presumed that these different pathways give rise to the two principal products of DMS oxidation, MSA and SO_2 (which is subsequently converted to sulphate). The decrease in R during warming might therefore reflect either increasing temperature in the DMS source region, or a change in the primary DMS source region, with increasing input from long-distance transport from the lower latitudes. Here, we make the implicit assumption that the basic features of the aerosol size distributions of MSA and n.s.s. sulphate (and hence their transport and depositional characteristics) are relatively constant for different climates. This point may bear further investigation in view of recent data indicating that the size distribution of MSA may be dependent on that of the preexisting aerosols¹³.

Other possible explanations for the generally low ratios observed in the high Antarctic plateau (for both glacial and interglacial conditions) require the input of a significant non-marine sulphate component. Stratospheric input has historically been considered a possible contribution to the sulphate budget of central Antarctica, but there is now general agreement that this contribution is weak except after sporadic major volcanic eruptions^{14–16}. Input of continental-derived sulphate is another possibility, as mentioned above. A significant component of the sulphate present during glacial times, at least, is ionically balanced by calcium and the possibility exists that this sulphate is terrestrially derived. During glacial stages (B, C and D) calcium concentrations in our samples vary widely (more than a factor of three, from 13 to 52 ng g^{-1}), but the ratio R remains essentially constant. This implies that this sulphate was marine in origin and was neutralized in the atmosphere by reaction with calcium carbonate of terrestrial origin.

Because the partitioning between MSA and n.s.s. sulphate seems to be modulated by climate, it is useful to examine the total oxidized sulphur, which we presume to be the sum of MSA and n.s.s. sulphate (Table 2) instead of n.s.s. sulphate alone as in previous studies^{5,6}. The higher concentrations of n.s.s. sulphate and MSA found in cold stages are partly due to the reduced accumulation rates during this time. To obtain a more representative picture of the actual atmospheric concentrations in the past, we have corrected concentrations in ice using a

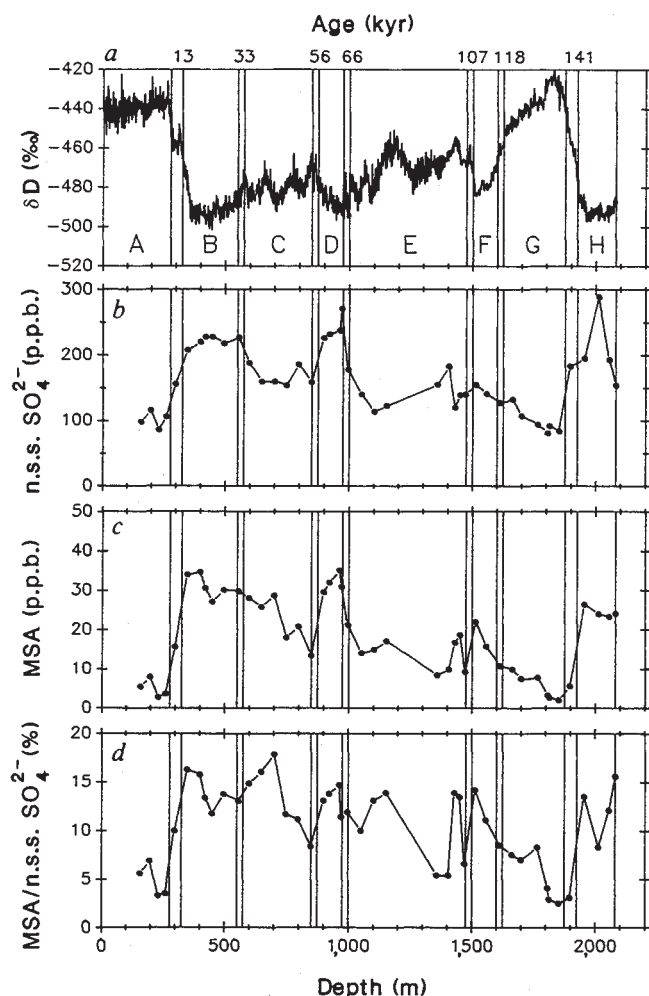


FIG. 1 Depth profiles of a, δD , b, non-seasalt sulphate, c, MSA and d, the ratio of MSA to non-seasalt sulphate in the Vostok ice core. Upper and lower stage boundaries are marked by vertical lines. (δD is deuterium isotope profile from ref. 3.)

method presented elsewhere⁶. We have also assumed that MSA and n.s.s. sulphate have similar deposition processes because both MSA and n.s.s. sulphate are present in the atmosphere as fine particles¹⁰. The corrected data, shown in Table 2, suggest that the oxidation products of DMS were constant during interglacial stages (stage A and G), but were enhanced significantly during the second part of the glacial age (stage B, C and D). Assuming that inputs of MSA and n.s.s. sulphate over the high Antarctic plateau are related largely to marine DMS emissions at mid-latitudes, our data suggest enhanced productivity from the DMS-producing portion of the oceanic biota between 18 and 70 kyr BP (before present). Whether this increase represents a real increase in oceanic productivity or simply an ecological shift favouring DMS producers is not known. Some workers

have suggested that marine sediment cores show evidence for increased productivity between 15 and 70 kyr BP in upwelling areas¹⁷, but this finding remains controversial.

Clearly, the ocean-atmosphere sulphur cycle exhibits variability that is intimately related to global climate change. It has been suggested that the biogenic aerosols themselves play an important part in forcing climate change either through direct radiative effects or through changes in cloud albedo and distribution^{6,18,19}. It may be possible to detect the effects of such climate forcing using high-resolution ice-core records which document the timing of changes in biogenic sulphur levels relative to those in stable isotopes, CO₂ and other chemical tracers during periods of rapid climate change. These studies are currently in progress. □

Received 19 September 1990; accepted 8 January 1991.

- Saigne, C. & Legrand, M. *Nature* **330**, 240–242 (1987).
- Legrand, M., De Angelis, M. & Delmas, R. *Analytica chim. Acta* **156**, 181–192 (1984).
- Jouzel, J. et al. *Nature* **329**, 403–408 (1987).
- Saigne, C., Kirchner, S. & Legrand, M. *Analytica chim. Acta* **203**, 11–21 (1987).
- Legrand, M., Lorius, C., Barkov, N. I. & Petrov, V. N. *Atmos. Envir.* **18**, 1867–1874 (1987).
- Legrand, M., Delmas, R. J. & Charlson, R. J. *Nature* **334**, 418–419 (1988).
- Legrand, M. & Delmas, R. J. *Nature* **327**, 671–676 (1987).
- Martin, J. H. & Fitzwater, S. E. *Nature* **331**, 341–343 (1988).
- Duce, R. A. *The Role of Air-Sea Exchange in Geochemical Cycling* (ed. Buat-Menard, P.) 497–529 (Reidel, Dordrecht, 1986).
- Saltzman, E. S., Savoie, D. L., Zika, R. G. & Prospero, J. M. *J. geophys. Res.* **88**, 10897–10902 (1983).
- Berresheim, H. *J. geophys. Res.* **92**, 13245–13262 (1987).

- Hynes, A. J., Wine, P. H. & Semmes, D. H. *J. phys. Chem.* **90**, 4149–4156 (1986).
- Pszenny, A. A. *J. atmos. Chem.* (submitted).
- Delmas, R. J. *Nature* **299**, 677–678 (1982).
- Cunningham, W. C. & Zoller, W. H. *J. Aerosol Sci.* **12**, 367–384 (1981).
- Legrand, M. & Delmas, R. J. *Atmos. Envir.* **18**, 1867–1874 (1984).
- Sarnthein, M., Winn, K. & Zahn, R. *Abrupt Climatic Change* (eds Berger, W. H. & Labeyrie, L. D.) 311–334 (Reidel, Dordrecht, 1987).
- Shaw, G. E. *Clim. Change* **5**, 297–303 (1983).
- Charlson, R. J., Lovelock, J. E., Andreae, M. O. & Warren, S. G. *Nature* **326**, 655–661 (1987).

ACKNOWLEDGEMENTS. Financial support was provided by the Centre National de la Recherche Scientifique and by the NSF. The Vostok ice core was analysed in the framework of a US–France–USSR cooperative agreement.

Amelioration of subsoil acidity by application of a coal-derived calcium fulvate to the soil surface

H. v. H. van der Watt*, R. O. Barnard*, I. J. Cronje†, J. Dekker†, G. J. B. Croft* & Maria M. van der Walt*

* Department of Soil Science and Plant Nutrition, University of Pretoria, 0002 Pretoria, South Africa

† Division of Energy Technology, CSIR, PO Box 375, 0001 Pretoria, South Africa

SUBSOIL acidity is a serious problem in many tropical and sub-tropical soils^{1–3}. The high acidity, low calcium contents and often toxic levels of soluble and/or exchangeable aluminium severely impair plant-root development in these soils^{1–3}. The relative immobility of surface-applied liming materials limits their ability to reduce subsoil acidity. Recently, the use of gypsum or phosphogypsum has been advocated as an alternative to lime^{3–7}. On the other hand, it is well known that humic substances can mobilize and form complexes with metals in soils^{8–11}. Here we report that a newly available, coal-derived calcium-fulvate is highly efficient as a carrier of calcium in the soil profile. Moreover, subsoil pH was considerably higher when calcium-fulvate was applied to the soil surface, than when gypsum, calcium-EDTA, Ca(OH)₂ or CaCO₃ were applied.

Fulvic acid was obtained from the wet oxidation of coal at elevated pressure and temperature, using a recently patented process¹². This process leads to the formation of a water-insoluble product (oxicoal, containing ~80% base-soluble 'humic' acids) and water-soluble 'fulvic' acids (10–15% yield). The fulvic acids, separated by filtration from the oxicoal, were neutralized to pH 7.2 with Ca(OH)₂.

We called the resulting 'solution' (actually a stable suspension because colloidal particles are usually still present) Ca-fulvate. The suspension could be dried to a fine, dark powder but, in our experiments, it was used without further drying. It contained 4.18% total solids and 11.26 g Ca l⁻¹.

TABLE 1 Relevant properties of experimental soils

Soil	Clay (<2 µm) (%)	pH (H ₂ O)	pH (1 mol l ⁻¹ KCl)	CEC* (soil) (cmol _c kg ⁻¹)	Exch.† bases (% of CEC)	Exch.† Ca (% of CEC)
Avalon	6.0	4.68	3.74	1.7	29.0	11.5
Hutton UP	20.2	5.02	4.08	3.2	39.7	18.8
Hutton PR	35.6	4.70	3.97	4.3	28.3	18.0

* Cation exchange capacity (cmol of charge per kg soil).

† Exchangeable.

The downward movement of Ca in soil was studied using cylinders 1 m long × 68 mm inside diameter. Plastic sheeting was used as lining to allow the easy removal of the entire soil column for segmentation and analysis. Care was taken to eliminate micro-channels which may result in preferential flow paths at the soil-liner boundary. Three soil types were used in our experiments (Table 1). The columns were carefully packed with dry soil and leached with two pore volumes of deionized water as a pretreatment to remove free salts. The following substances were then added in 150-cm³ suspensions to the surface of the water-saturated soil column: H₂O (control), CaCO₃, Ca(OH)₂ (freshly prepared), CaSO₄·2H₂O, Ca-EDTA and Ca-fulvate. In the Ca treatments 1.689 g Ca was added in each case; this is equivalent to 20 t ha⁻¹ gypsum or 11.6 t ha⁻¹ CaCO₃. The columns were leached with deionized water at a flow rate of 55 mm h⁻¹ until two pore volumes of leachate had been collected in measured fractions. This flow rate is fairly high, but of course lower than the hydraulic conductivity of the columns. The effect of flow rate on Ca distribution may be important, but was not studied. After cessation of water inflow, the columns were allowed to drain. The drainage water was collected as a final leachate fraction. The soil column encased within the lining could easily be removed without disturbance, and was cut into 100-mm-long segments. The segments were dried at 353 K and weighed. The leachate fractions were analysed for Ca, Mg, Na, K, Fe and Al and the pH determined. The soil segments were analysed for the following: total soluble and exchangeable Ca, Mg, Na and K extractable with 1 mol l⁻¹ ammonium acetate (pH 7) using a 1:10 soil/solution ratio; soluble cations in a 1:1

TABLE 2 Calcium derived from surface-applied suspensions and leached to below 150 mm in soil columns

Soil and treatment	Ca in excess of control found in:		Percentage of applied Ca leached to:		
	150-430 mm (mg)	430-920 mm (mg)	leachate (mg)	>150 mm (%)	>430 mm (%)
Avalon					
CaSO ₄	112	147	205	27.5	20.8
Ca-EDTA	183	190	5	22.4	11.5
Ca-fulvate	339	384	375	65.0	44.9
Hutton UP					
CaSO ₄	147	0	23	10.1	1.4
Ca-EDTA	321	234	20	34.0	15.0
Ca-fulvate	352	368	268	58.5	37.7
Hutton PR					
CaSO ₄	183	62	0	14.5	3.7
Ca-EDTA	365	120	0	28.7	7.1
Ca-fulvate	293	267	189	44.3	27.0

soil/water extract; and Fe and Al extractable with 1 mol l⁻¹ KCl using a 1:10 soil/solution ratio. All cations were determined using standardized atomic-absorption spectrometry procedures.

Some of our results are shown in Figs 1 and 2. We consistently found that soil pH increased with depth where Ca-fulvate was applied, but when gypsum was applied, we always found that soil pH became lower in contrast to findings reported elsewhere³. The exchangeable-plus-soluble soil Ca at depths below 150 mm and 430 mm in the soil increased dramatically where Ca-fulvate was applied. This effect was much less marked for applications of Ca-EDTA and gypsum, whereas CaCO₃ and Ca(OH)₂ did not provide Ca to the deeper soil zones, as would be expected considering their low solubilities (the data for Ca(OH)₂ and CaCO₃ are not shown for the sake of brevity). In Table 2 the percentage of Ca derived from the surface-applied substances and leached beyond the above depths in the soil columns is shown for three of the Ca compounds. Where Ca-fulvate was used, substantial quantities of Ca (11-22% of the Ca applied) were found in the leachate. Because the term fulvic acid refers to a group of substances¹⁰ with molecular weights in the range 500-2,000, this result indicates that at least some fraction of the

TABLE 3 Total cations leached from 1-m long soil columns surface-treated with various Ca compounds for the Hutton soil

Compound added	Total cations leached					
	Ca	Mg	Na (mg)	K	Fe	Al
0	4	2	0	4	0	6
CaCO ₃	2	1	0	3	0	2
Ca(OH) ₂	7	2	0	8	0	2
CaSO ₄	3	2	1	4	1	2
Ca-EDTA	3	10	9	9	27	340
Ca-fulvate	193	78	8	15	1	37

Ca-fulvate complex is very stable and releases its Ca very slowly or not at all while moving down the soil column. This was confirmed by the pattern of the breakthrough curves, which show a sharp peak after about one pore volume of leachate has been collected following surface application of the Ca-fulvate suspension. Figure 2 shows clearly that Ca-fulvate is effective in supplying Ca to the subsoil.

The higher soil pH which results from the application of Ca-fulvate may be due to one or more of the following reactions: the release of some Ca to the soil solution through the hydrolysis of Ca-fulvate, which occurs continuously as the soluble Ca-fulvate is leached downward; the replacement of exchangeable Al and H by Ca, followed by the removal of Al and H, complexed with the fulvic acid, either by leaching to lower soil horizons or carried away in the leachate. The latter mechanism seems to apply particularly to EDTA, as considerable amounts of H, Al and Fe were found in the leachates from the EDTA treatments (Table 3). Leaching is an essential part of the process for any surface-applied amendment. Furthermore, reactive hydroxyls may also be carried by the fulvate (fulvic acid neutralized to pH 7.2), or as the ion pair Ca(OH)⁺, which would result in the direct neutralization of soil acidity. The differences in the results for Ca-EDTA and Ca-fulvate may in part be explained by the observation that, in the case of Ca-EDTA, much of the Ca is retained within the top 5 cm of soil (data not shown). The exchangeable acidity replaced by Ca in this way would explain the lowering of the pH deeper in the soil, whereas the more

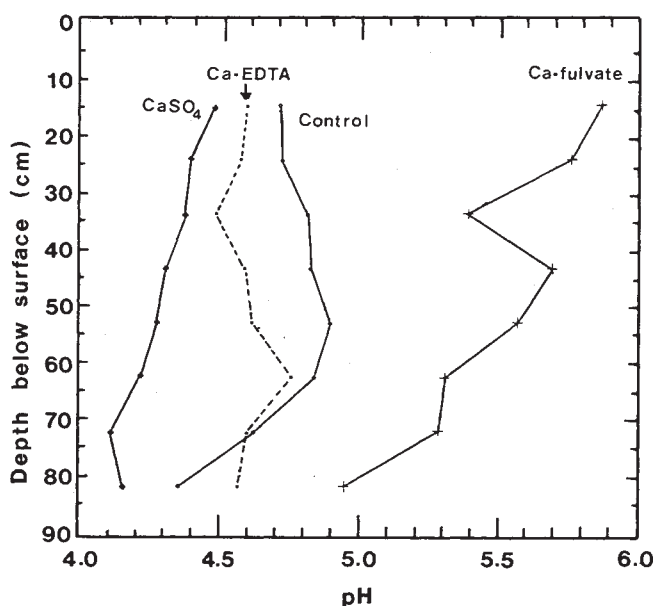
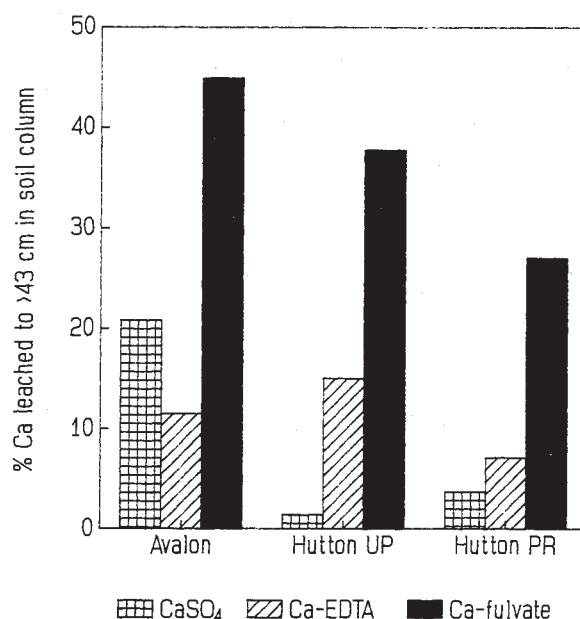
FIG. 1 Soil pH/depth relationships for the Avalon soil following surface applications of CaSO₄, Ca-EDTA and Ca-fulvate and leaching.

FIG. 2 Diagram showing the percentage of applied Ca from three surface-applied substances leached to a soil depth >43 cm, for three soils.

stable Al-EDTA appears in the leachate (Table 3). In contrast, owing possibly to differences in stability constants and/or reaction rates, the Ca from the Ca-fulvate is released more gradually, or is leached to below the 1 m depth (Table 3).

We conclude that the use of coal-derived Ca-fulvate as an ameliorant for acid and Ca-deficient subsoils is very promising, particularly as only moderate leaching was used in our experiments. But considerable further research is required on aspects such as product optimization and economics, effects on plants and possible groundwater contamination with metal-fulvate complexes. At present we estimate that the product will cost less than 300 dollars per tonne. In practice, use thereof will be determined by crop response and the relevant studies are under way. □

Received 11 June; accepted 31 December 1990.

1. Bouldin, D. R. *Cornell Int. Agric. Bull.* No. 34 (1979).
2. Haynes, R. J. *Adv. Agron.* **37**, 249–315 (1984).
3. Shainberg, I. et al. *Adv. Soil Sci.* **9**, 1–111 (1989).
4. Reeve, N. G. & Sumner, M. E. *Agrochimica* **4**, 1–6 (1972).
5. Farina, M. P. W. & Channon, P. *Soil Sci. Soc. Am. J.* **52**, 175–180 (1988).
6. Pavan, M. A., Bingham, F. T. & Pratt, P. F. *Soil Sci. Soc. Am. J.* **46**, 1201–1207 (1982).
7. Oates, K. M. & Caldwell, A. G. *Soil Sci. Soc. Am. J.* **49**, 915–918 (1985).
8. Gregor, J. E., Powell, H. K. T. & Town, R. M. *J. Soil Sci.* **40**, 661–673 (1989).
9. Saar, R. A. & Weber, J. H. *Environ. Sci. Tech.* **16**, 510A–517A (1982).
10. Stevenson, F. J. *Humus Chemistry: Genesis, Composition Reactions* (Wiley, New York, 1982).
11. Aiken, R. A., McKnight, D. M., Wershaw, R. L. & MacCarthy, P. (eds) *Humic Substances in Soil, Sediment, and Water* (Wiley, New York, 1985).
12. Cronje, I. J. *Process for the Oxidation of Coal*. S. Afr. Patent 88/4770 (1988).
13. Lindsay, W. L. *Chemical Equilibria in Soils* (Wiley, New York, 1979).

Cryptomonad algae are evolutionary chimaeras of two phylogenetically distinct unicellular eukaryotes

Susan E. Douglas*, Colleen A. Murphy*,
David F. Spencer† & Michael W. Gray†

* Institute of Marine Biosciences, National Research Council of Canada,
1411 Oxford Street, Halifax, Nova Scotia B3H 3Z1, Canada

† Department of Biochemistry, Dalhousie University, Halifax,
Nova Scotia B3H 4H7, Canada

ALTHOUGH it is widely accepted that the plastids of plants and algae originated as endosymbionts¹, the details of this evolutionary process are unclear^{2,3}. It has been proposed that in organisms whose plastids are surrounded by more than two membranes, the endosymbiont was a eukaryotic alga rather than a photosynthetic prokaryote⁴. The DNA-containing⁵ nucleomorph⁶ of cryptomonad algae appears to be the vestigial nucleus of such an algal endosymbiont⁷. Eukaryotic-type ribosomal RNA sequences have been localized to a nucleolus-like structure in the nucleomorph⁸. In support of the hypothesis that cryptomonads are evolutionary chimaeras of two distinct eukaryotic cells, we show here that *Cryptomonas* Φ contains two phylogenetically separate, nuclear-type small-subunit rRNA genes, both of which are transcriptionally active. We incorporate our rRNA sequence data into phylogenetic trees, from which we infer the evolutionary ancestry of the host and symbiont components of *Cryptomonas* Φ. Such trees do not support the thesis³ that chromophyte algae evolved directly from a cryptomonad-like ancestor.

To identify nuclear-type small-subunit (18S) rRNA genes in a cryptomonad alga, we amplified total *Cryptomonas* Φ DNA by polymerase chain reaction (PCR) using primers specific for nuclear-encoded 18S rRNA sequences⁹. This gave two products, one (Nu) of 1.8 kilobase pairs (kb) and another (Nm) ~200 base pairs (bp) longer. The sequence of the Nu amplification product corresponds to that of the 18S rRNA gene encoded by

a 9.5-kb nuclear rDNA repeat, independently cloned from the main DNA band of a Hoechst dye-CsCl density gradient and completely sequenced (D.F.S., unpublished results). The Nm product represents a distinct, novel 18S rRNA gene, which we tentatively assign to the nucleomorph. When applied to total DNA from non-cryptomonad algae, PCR amplification yields only a single product, as expected.

The Nu and Nm sequences (1,773 and 2,039 nucleotides, respectively) are aligned in Fig. 1. Although the primary sequences are very similar (85.2%, excluding variable regions¹⁰ V2 and V4), it is evident that the Nm gene is not simply a variant of the Nu gene, but is structurally distinct. Secondary-structure models (data not shown) demonstrate that both genes encode potentially functional 18S rRNAs, with variable regions confined to portions of the secondary structure previously recognized as divergent among small-subunit rRNAs¹¹. There are no obvious structural aberrations to suggest that the Nm gene is an inactive pseudogene; on the contrary, the Nm and Nu secondary structures both contain the complete conserved core¹¹ that is critical for small-subunit rRNA function¹².

Using the Nm clone to probe Southern blots of *Cryptomonas* total cellular DNA, we found that the Nm and Nu rRNA genes originate from DNA components of widely different base composition (Fig. 2). The Nm probe hybridized to a 7.5-kb *Pst*I restriction fragment that is enriched in a high-(A+T) fraction (fraction 2, Fig. 2) of *Cryptomonas* DNA; this fragment is distinct from the main *Pst*I fragment (>45 kb), tentatively assigned to mitochondrial DNA¹³, also found in this region of a Hoechst dye-CsCl density gradient. Cross-hybridization of the Nm probe to the nuclear rDNA repeat is evident in fractions 7–9. The Nm probe did not hybridize to plastid rDNA-containing *Pst*I fragments of 8.6 and 10.0 kb¹³ in fractions 4 and 5.

To determine whether the Nm and Nu genes are both expressed, complementary oligodeoxynucleotide probes specific for each sequence (A, B and C; Fig. 1) were hybridized to duplicate northern blots. Probe A (Nu-specific) hybridized to a prominent RNA band of ~1.8 kb, in agreement with the size predicted from the DNA sequence (lane 2, Fig. 3). Probe B (Nm-specific) hybridized to an RNA species of 1.45 kb, which is considerably smaller than expected from the gene sequence (lane 3, Fig. 3). A third probe (C, specific for the Nm 5'-terminal region) detected an additional small RNA of ~250 nucleotides (lane 4, Fig. 3). These results not only verify that the Nm gene is transcribed, but also indicate that the mature Nm small-subunit rRNA is discontinuous. Interestingly, the break coincides with an apparent 166-nucleotide insertion in helix E10–11 (ref. 10) contained within variable region V2; this insert must represent a site of post-transcriptional spacer excision. A discontinuity in the V2 region has been described in the case of the mitochondrial small-subunit rRNA of *Tetrahymena pyriformis*¹⁴, but the split Nm rRNA reported here is the first example of a discontinuous nucleocytoplasmic 18S rRNA.

Although proof that the Nm rRNA gene is encoded by nucleomorph DNA must await isolation of a nucleomorph-enriched subcellular fraction and/or *in situ* hybridization experiments using Nm-specific probes, our data strongly support the idea that the cryptomonad nucleomorph contains active eukaryotic-type rRNA genes. Phylogenetic analysis by construction of rRNA trees using both maximum parsimony^{15,16} and neighbour-joining¹⁷ methods, shows that the *Cryptomonas* Nm and Nu sequences are evolutionarily distant from one another (Fig. 4). Regardless of the particular tree construction method used, three features of these nuclear small-subunit rRNA trees are consistently observed. (1) The Nm sequence branches as a specific relative of two rhodophyte nuclear 18S rRNA sequences, supporting the hypothesis that the cryptomonad endosymbiont was a primitive red alga⁴; (2) the Nu sequence branches separately from the Nm sequence and in the same assemblage as the amoeboid protozoan *Acanthamoeba castellanii* (in the tree

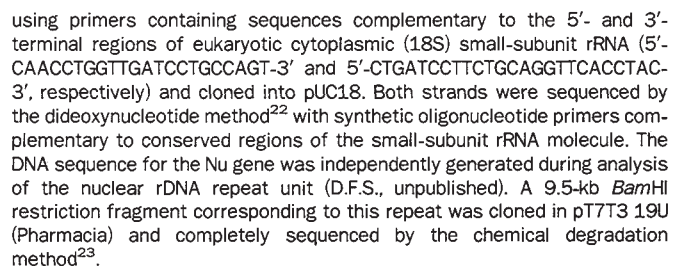


FIG. 1. Alignment of the nuclear (Nu) and putative nucleomorph (Nm) small-subunit rRNA sequences of *Cryptomonas* Φ . The alignment is based on secondary-structure models constructed for the two (data not shown). Dashes indicate apparent deletions and asterisks denote nucleotide identities shared by the two sequences. Principal variable regions (V2, V4)¹⁰ are enclosed by solid lines. Sequences corresponding to the universally conserved regions (U1–U8)¹¹ that define the functional small-subunit rRNA core are underlined. Positions of complementary oligodeoxynucleotide probes (A, B, C) used in northern hybridizations (see Fig. 3) are indicated by overlining. METHODS. Total cellular DNA was prepared from *Cryptomonas* Φ (ref. 13) and used as a substrate for PCR. Ribosomal RNA genes were amplified⁹

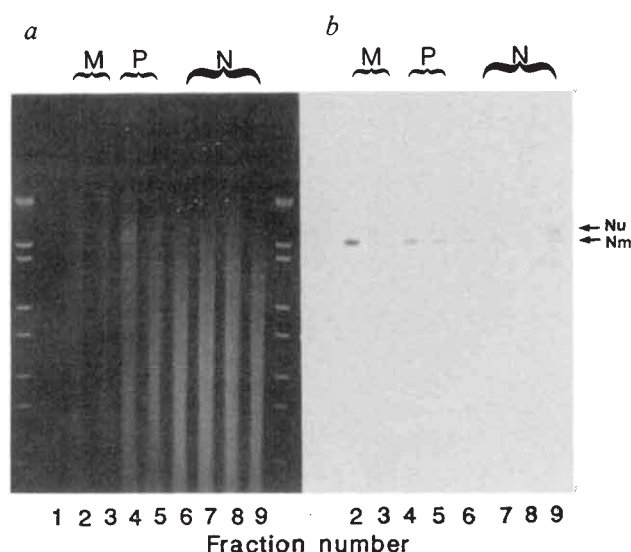
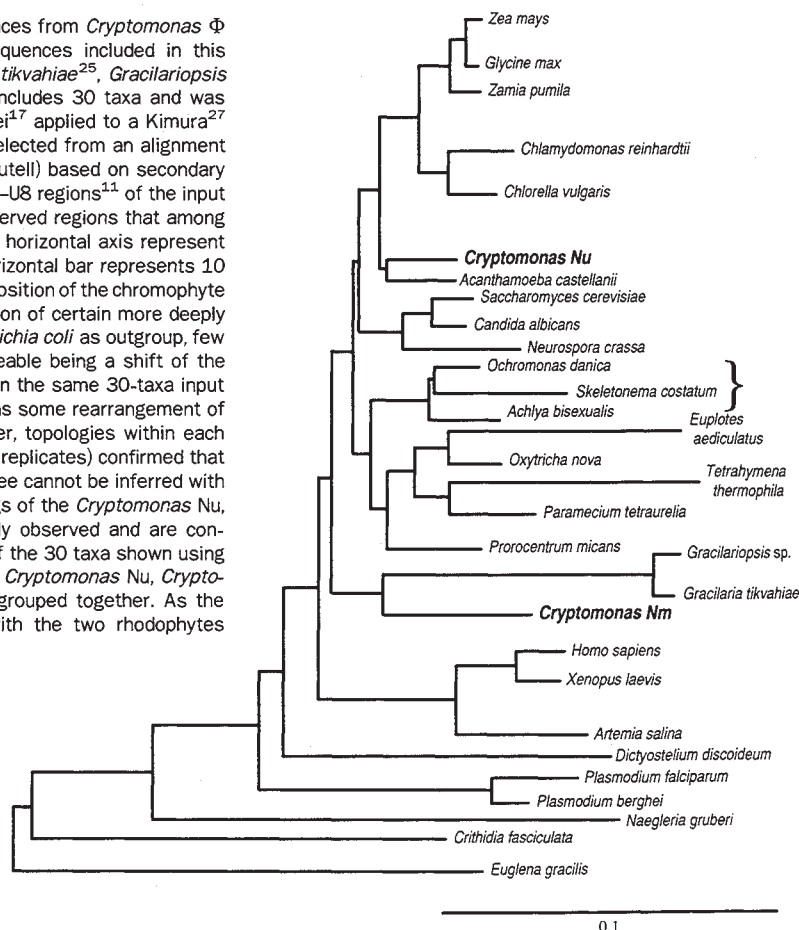


FIG. 2 Detection of rRNA coding sequences in the total DNA from *Cryptomonas* Φ . *a*, Agarose-gel electrophoresis of total DNA fractionated in a Hoechst dye–CsCl density gradient and digested with *Pst*I. Fractions containing mitochondrial (M), plastid (P) and nuclear (N) DNA are indicated. Markers are *Sty*I restriction fragments of bacteriophage λ DNA (24.5, 7.7, 6.2, 3.4, 2.7, 1.89, 1.45 kb). *b*, Autoradiogram of Southern blot hybridized with the Nm clone. The bands corresponding to the nuclear (Nu) and nucleomorph (Nm) rDNA-containing fragments are indicated by arrows. Restriction endonuclease analysis of aliquots from Hoechst 33258 dye–CsCl density gradients, transfer to Zeta Probe membranes (BioRad) and Southern hybridization analysis were as described previously¹³.

FIG. 4 Phylogenetic analysis of small-subunit rRNA coding sequences from *Cryptomonas* Φ and other eukaryotes. Nuclear-encoded small-subunit rRNA sequences included in this analysis are from ref. 10, with the exception of those from *G. tikvahiae*²⁵, *Gracilariopsis* sp.²⁵ and *Candida albicans*²⁶. This particular phylogenetic tree includes 30 taxa and was constructed using the neighbour-joining method of Saitou and Nei¹⁷ applied to a Kimura²⁷ dissimilarity matrix. The data set consisted of 1,001 positions selected from an alignment of eukaryotic small-subunit rRNA sequences (supplied by R. R. Gutell) based on secondary structure comparisons; in addition to the universally conserved U1–U8 regions¹¹ of the input small-subunit rRNAs, the database encompasses other well conserved regions that among eukaryotes can be aligned unambiguously. Branch lengths on the horizontal axis represent the evolutionary distance between the nodes of the tree. The horizontal bar represents 10 changes per 100 nucleotides. The bracket indicates the branching position of the chromophyte algae. When the number of taxa was expanded to 47, with inclusion of certain more deeply branching eukaryotes (*Vairimorpha*, *Giardia*) and employing *Escherichia coli* as outgroup, few changes in overall tree topology were observed, the most noticeable being a shift of the two *Plasmodium* sequences to a position higher in the tree. When the same 30-taxa input data were analysed by a maximum parsimony method¹⁶, there was some rearrangement of major branches relative to the topology shown in Fig. 4; however, topologies within each major group were identical or very similar. Bootstrap analysis (20 replicates) confirmed that the specific branching orders for the major groups higher in the tree cannot be inferred with high confidence. But the relative positions and separate branchings of the *Cryptomonas* Nu, *Cryptomonas* Nm and chromophyte sequences were consistently observed and are considered robust. These branchings are supported by the analysis of the 30 taxa shown using the DNAMove program¹⁶ and starting with an initial tree in which *Cryptomonas* Nu, *Cryptomonas* Nm and the chromophyte *Skeletonema costatum* were grouped together. As the analysis proceeded, *Cryptomonas* Nm affiliated immediately with the two rhodophytes (*Gracilariopsis* sp. and *Gracilaria tikvahiae*), while *Cryptomonas* Nu was progressively displaced away from the chromophytes toward the fungal/chlorophyte/land plant grouping.



shown in Fig. 4, the two form a sister group to chlorophytes and land plants); (3) the *Cryptomonas* Nu sequence does not branch together with members of the chromophyte algae (bracketed arrow, Fig. 4).

Ribosomal RNA sequence comparisons have supported the view that multiple endosymbioses occurred in the formation of land plants and algae¹⁸⁻²⁰. It has been argued³ that only two endosymbiotic events need be invoked to explain the diversity of extant photosynthetic eukaryotes, the first (phagotrophic eukaryotic host plus cyanobacterium) producing those organisms (rhodophytes, chlorophytes and land plants) whose plastids are surrounded by two membranes, the second (a different phagotrophic eukaryote plus primitive eukaryotic alga) giving rise to those algae (including cryptophytes and chromophytes) whose plastids are surrounded by more than two membranes. The latter grouping comprises the proposed kingdom Chromista²¹. Because it is postulated that chromophytes arose from cryptophytes by loss of phycobiliproteins and evolution of chlorophyll *c*₁ (refs 3, 21), cryptomonads occupy a pivotal position in defining the Chromista. In our analysis, chromophytes form a group separate from *Cryptomonas*, which does not support the concept (based largely on ultrastructural data) of a monophyletic kingdom Chromista²¹, in which chromophyte algae are directly derived from a cryptomonad-like ancestor. Our study indicates that cryptophytes, chlorophytes and land plants shared a common ancestor distinct from that which gave rise to the chromophytes. This conclusion is further supported by comparisons of large-subunit²⁰ and 5S (ref. 5) rRNA sequences and by morphological data⁵, all of which suggest that cryptomonads are not particularly closely related to any other algal group.

In conclusion, our results strengthen the hypothesis that the evolutionary formation of cryptomonad algae progressed through two separate stages: a primary endosymbiotic event involving a eukaryotic host and a photosynthetic prokaryote; and a secondary event involving the resulting eukaryotic alga as endosymbiont and a second eukaryote as host. From the small-subunit rRNA sequences encoded by the nuclear genomes of the two eukaryotes contributing to the cryptomonad chimera, we have been able to discern aspects of the evolutionary history of this complex organism and to assess proposed endosymbiotic events which gave rise to the variety of photosynthetic eukaryotes seen today. Although rRNA sequence comparisons should not be regarded as the only means by which phylogenies can be inferred, a re-evaluation of the concept of the kingdom Chromista²¹ as a monophyletic assemblage would be in order, on the basis of our proposal that chromophyte and cryptophyte algae probably arose by different endosymbiotic events. □

Received 17 August; accepted 12 December 1990.

- Gray, M. W. *Trends Genet.* **5**, 294-299 (1989).
- Whitley, J. M. & Whitley, F. R. *New Phytol.* **87**, 233-247 (1981).
- Cavallier-Smith, T. *Biol. J. Linn. Soc.* **17**, 289-306 (1982).
- Gibbs, S. P. *Ann. N.Y. Acad. Sci.* **391**, 193-208 (1981).
- Ludwig, M. & Gibbs, S. P. *Protoplasma* **127**, 9-20 (1985).
- Greenwood, A. D., Griffiths, H. B. & Santore, U. J. *Br. Phycol. J.* **12**, 119 (1977).
- Ludwig, M. & Gibbs, S. P. *Ann. N.Y. Acad. Sci.* **503**, 198-211 (1987).
- McFadden, G. I. *J. Cell Sci.* **95**, 303-308 (1990).
- Sogin, M. L. in *PCR Protocols. A Guide to Methods and Applications* (eds Innes, M. A. et al.) 307-314 (Academic, San Diego, 1990).
- Neefs, J.-M., Van de Peer, Y., Henriks, L. & De Wachter, R. *Nucleic Acids Res.* **18** (suppl.), 2237-2247 (1990).
- Gray, M. W., Sankoff, D. & Cedergren, R. *J. Nucleic Acids Res.* **12**, 5837-5852 (1984).
- Dahlberg, A. E. *Cell* **57**, 525-529 (1989).
- Douglas, S. E. *Curr. Genet.* **14**, 591-598 (1988).
- Schnare, M. N., Heinonen, T. Y. K., Young, P. G. & Gray, M. W. *J. Biol. Chem.* **261**, 5187-5193 (1986).
- Swofford, D. L. PAUP Version 3.0. Illinois Natural History Survey (Champaign, Illinois, 1989).
- Felsenstein, J. PHYLIP Manual Version 3.3 (Herbarium, Univ. California, Berkeley, 1990).
- Saitou, N. & Nei, M. *Molec. Biol. Evol.* **4**, 406-425 (1987).
- Gunderson, J. H., Elwood, H. J., Ingold, A., Kindle, K. & Sogin, M. L. *Proc. natn. Acad. Sci. U.S.A.* **84**, 5823-5827 (1987).
- Bhattacharya, D., Elwood, H. J., Goff, L. J. & Sogin, M. L. *J. Phycol.* **26**, 181-186 (1990).
- Perasso, R., Baroin, A., Qu, L. H., Bachellerie, J. P. & Adoutte, A. *Nature* **339**, 142-144 (1989).
- Cavallier-Smith, T. *Prog. Phycol. Res.* **4**, 309-347 (1986).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499-560 (1980).
- Douglas, S. E. & Durnford, D. G. *Plant molec. Biol.* **13**, 13-20 (1989).

- Bird, C. J., Rice, E. L., Murphy, C. A., Liu, Q. Y. & Ragan, M. A. *Nucleic Acids Res.* **18**, 4023-4024 (1990).
- Hendriks, L. et al. *System appl. Microbiol.* **12**, 223-229 (1989).
- Kimura, M. *J. molec. Evol.* **16**, 111-120 (1980).

ACKNOWLEDGEMENTS. We thank J. Gallant for technical assistance, M. Ragan for critical review, and R. R. Gutell for alignment of eukaryotic 18S rRNA sequences, from which the database used for phylogenetic tree construction was selected. This work was supported by an operating grant from the Medical Research Council of Canada to M.W.G., who is a Fellow in the Evolutionary Biology Program of the Canadian Institute for Advanced Research. This is NRCC publication number 31933. The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers X57162 (nuclear) and X57008 (nucleomorph).

Worldwide migration of amplified insecticide resistance genes in mosquitoes

Michel Raymond*, Amanda Callaghan*†, Philippe Fort‡ & Nicole Pasteur*

* Laboratoire Génétique et Environnement, Institut des Sciences de l'Évolution, URA 327, Case Courrier 64; and † Laboratoire de Biologie Moléculaire, URA 1191, Case Courrier 12; Université Montpellier II, 34095 Montpellier, France

IN *Culex pipiens*, overproduction of nonspecific esterases is a common mechanism of resistance to organophosphate insecticides^{1,2}. The esterases are attributed to closely linked loci named A and B according to substrate preference³⁻⁶, and overproduction of all esterases B is due to gene amplification^{7,8}. Distribution of electrophoretically distinct variants of overproduced esterases A and B is geographically restricted, with the exception of esterases A2 and B2, always found together throughout at least three continents (Fig. 1). To determine whether this situation is due to migration or to a high mutation rate, esterase B structural genes and their flanking regions were compared by sequence and/or restriction fragment length polymorphism analysis. Whereas structural genes were similar, flanking regions of electrophoretically dissimilar esterases B varied considerably. In contrast, flanking sequences of esterases B2 from different geographical locations (Africa, Asia, North America) were identical. These results suggest that amplified esterase B2 genes originated from an initial event that has subsequently spread organophosphate insecticide resistance by migration.

To study the genetic relationship between esterases B1 and B2, the *HindIII*-*Bam*HI fragment of the esterase B2 gene was cloned, sequenced and compared with the homologous esterase B1 fragment⁹. The two showed 96% nucleotide sequence homology (not shown), and 97% homology in the predicted amino-acid sequence (Fig. 2). Both esterase B1 and B2 clones were used as probes to construct restriction maps of esterases B from two strains, S-Lab and MSE, which contain a single copy of the gene (each producing an esterase B with a particular electrophoretic mobility different from that of esterase B1 and B2), as well as from mosquitoes with an amplified esterase B1 (Tem-R strain) or an amplified esterase B2 (six strains from different continents: Africa, North America and Asia). Under stringent detection conditions, each map was identical when constructed using either esterase B1 or B2 probes, demonstrating that the probes recognize exactly the same DNA regions. These two results strongly suggest that before amplification, esterase B1 and B2 genes were allelic. In addition, they show that esterase B1 and B2 genes were also alleles of the unamplified esterase B genes observed in S-Lab and MSE. This agrees with results from crossing experiments¹⁰ which established that various unamplified esterases B (different from B1 and B2) are alleles.

A comparison between the different restriction maps (Fig. 3) indicates that the DNA sequences are very similar within the

† Present address: School of Animal and Microbial Science, University of Reading, Whiteknights, PO Box 228, Reading, RG6 2AJ, UK.

gene among mosquitoes with different esterases. But the number of conserved restriction sites in the sequences flanking the structural genes is relatively small, indicating that the number of mutation events separating any two sequences flanking an esterase B gene is high and that their genetic divergence is probably ancient. In contrast, the whole of the DNA sequences,

analysed by restriction mapping, from strains with esterase B2, collected from regions covering three continents, were identical.

The level of polymorphism of the unamplified B locus and its flanking sequences in single mosquitoes was also very high. After only a *Hind*III digestion, many alleles were detected (data not shown): at least 15 in a sample of 30 insects from Saint-

FIG. 1 Known geographic distribution of overproduced esterases B1 (○) and A2-B2 (▲) associated with resistance to organophosphates in mosquitoes of the *Culex pipiens* complex^{3,4,11-13,21-25}. The distribution of esterases A2-B2 has increased over the last decade. For example, in California, they appeared before 1984 in an area where until 1980 only esterase B1 had been found¹²; in southern France they were detected in 1986¹³ in regions where previously only esterase A1 was present²⁶, and in Italy they were absent in 1982²⁷ but present in 1985³.

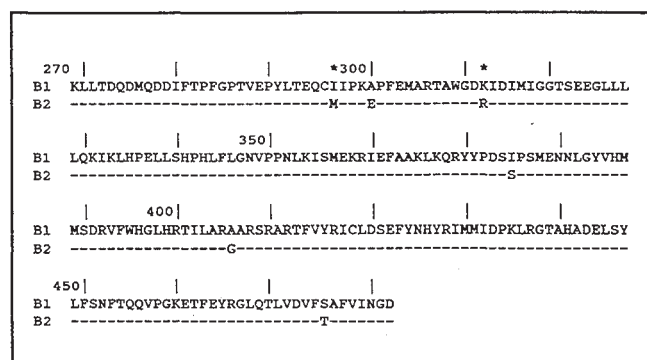
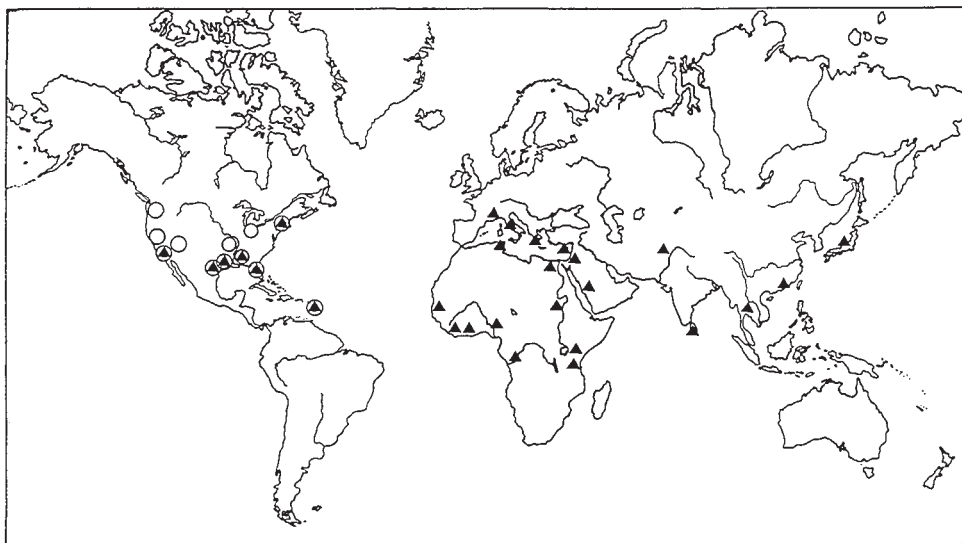
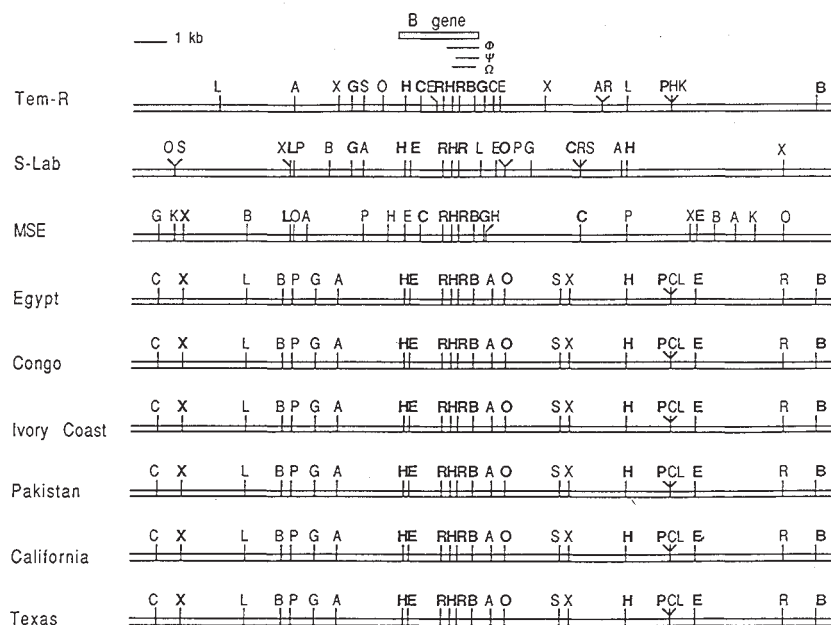


FIG. 2 Predicted amino-acid sequence similarity (single-letter notation) between esterase B1 and B2 gene products (amino acids 269-482). A fragment of the esterase B2 gene was isolated from a Puc-19 derived *Bam*HI-*Hind*III genomic library from the SeLax strain (homozygous for the presence of amplified esterase B2). DNA sequencing used M13 subcloning and dideoxy sequencing²⁸. The esterase B2 fragment sequence was compared to that of the esterase B1 *Hind*III-*Bam*HI fragment⁹. Conservative amino acids are indicated by an asterisk. The esterase B1 protein sequence is numbered as previously published.

FIG. 3 Restriction maps of the esterase B genes from organophosphate insecticide-resistant and susceptible adult mosquitoes. The strains used were: Tem-R (ref. 29) from California containing ~250 copies of the esterase B1 gene; S-Lab (ref. 30) a susceptible strain from California; MSE (ref. 31) a low esterase activity strain from southern France; and six strains containing amplified esterase B2 (amplification level ~60-fold), namely Houston²⁴ and SeLax¹² (Texas and California, North America), Lahore²⁴ (Pakistan, Asia), and Bouaké¹³, Brazaville and Soft³² (Ivory Coast, Congo and Egypt, Africa).

METHODS. Genomic DNA was extracted from adult mosquitoes using a phenol-chloroform method⁸, and digested with one or several restriction enzymes: *Acc*I (A), *Bam*HI (B), *Bcl*I (C), *Bgl*II (G), *Hind*III (H), *Eco*RI (E), *Eco*RV (R), *Kpn*I (K), *Pst*I (P), *Sac*I (S), *Sal*I (L), *Xba*I (X) and *Xho*I (O). The resulting DNA fragments were blotted onto nylon filters (Hybond^N) after the method of Southern³³, and visualized by hybridization to the ³²P-labelled 0.7- or 1.3-kilobase (kb) complementary DNA clones of the esterase B1 gene^{7,9}, or to the 0.7-kb genomic esterase B2 clone. The filters were washed stringently to ensure that only closely related DNA fragments were being studied. Above the restriction maps are indicated the localizations of the esterase B gene (shaded box), of the 1.3-kb (Φ) and 0.7-kb (Ψ) cDNA esterase B1 probes, and of the 0.7-kb (Ω) esterase B2 genomic probe.



Martin-de-Londres (France), and 13 (all different from the first 15) in a sample of 28 mosquitoes from Sintra (Portugal). None of these were unamplified B1 or B2 alleles.

These observations suggest that all amplified esterase B2 genes and flanking sequences have arisen from a single initial event and support the hypothesis that the worldwide presence of esterase B2 amplification is due to recent migration. This conclusion agrees well with the strong linkage disequilibrium, also observed worldwide, between esterases B2 and A2, which can best be explained by a rapid migration of an initial A2-B2 gene association. The place and time of the occurrence of the original amplification event cannot be traced, but Africa or Asia is most likely^{4,11}. The recent history of the sudden appearance of esterases A2-B2 is well documented for California, Italy and France^{3,12,13} and a migration between continents is not unlikely, as live larvae and adults of the *C. pipiens* complex are often found on international flights and boats¹⁴⁻¹⁶.

The rapid spread of B2 esterase gene amplification since its probable occurrence in the 1960s (when the intensive use of organophosphate insecticides began) can be compared with that of the *Drosophila melanogaster* alcohol dehydrogenase *Fast* allele associated with high alcohol dehydrogenase activity. There is less restriction polymorphism among *Fast* than *Slow* alleles¹⁷⁻²⁰. This has been interpreted to reflect a recent origin of the *Fast* allele, perhaps within the past 2,000 years¹⁷. Its presence, mainly in temperate areas, may result from an extensive migration and an unknown selective factor.

Resistance genes (or genetic factors such as gene amplification) may arise in a population under insecticide control by mutation or by migration. In *C. pipiens*, very few factors conferring significant organophosphate resistance have been selected during the past 30 years, indicating that the genetic possibilities are limited. As each has a relatively low probability of occurrence, the evolution of organophosphate resistance in a treated population is more likely to be determined by the extent of migration rather than by the rate of mutation. □

Received 16 August; accepted 21 December 1990.

1. Fournier, D. et al. *Pest. Biochem. Physiol.* **27**, 211-217 (1987).
2. Mouchès, C. et al. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2113-2116 (1987).
3. Callaghan, A. thesis, Univ. of London (1989).
4. Villani, F., White, G. B., Curtis, C. F. & Miles, S. J. *Bull. ent. Res.* **73**, 153-170 (1983).
5. Pasteur, N., Iseki, A. & Georgiou, G. P. *Biochem. Genet.* **19**, 909-919 (1981).
6. Wirth, M. C., Marquine, M., Georgiou, G. P. & Pasteur, N. *J. med. Entomol.* **27**, 202-206 (1990).
7. Mouchès, C. et al. *Science* **233**, 778-780 (1986).
8. Raymond, M. et al. *Biochem. Genet.* **27**, 417-423 (1989).
9. Mouchès, C. et al. *Proc. natn. Acad. Sci. U.S.A.* **87**, 2574-2578 (1990).
10. de Stordeur, E. *Biochem. Genet.* **14**, 481-493 (1976).
11. Curtis, C. F. & Pasteur, N. *Bull. ent. Res.* **71**, 153-161 (1981).
12. Raymond, M. et al. *J. med. Ent.* **24**, 24-27 (1987).
13. Magnin, M. thesis, Univ. Paris VI (1986).
14. Andreadis, T. G. *J. Am. Mosq. Control Assoc.* **4**, 256-260 (1988).
15. Asahina, S. *Jap. J. Med. Sci. Biol.* **23**, 255-258 (1970).
16. Highton, R. B. & Van Someren, E. C. C. *Bull. WHO* **42**, 334-335 (1970).
17. Aquadro, C. F., Desse, S. F., Bland, M. M., Langley, C. H. & Laurie-Ahlberg, C. C. *Genetics* **114**, 1165-1190 (1986).
18. Kreitman, M. & Aguadé, M. *Proc. natn. Acad. Sci. U.S.A.* **83**, 3562-3566 (1986).
19. Aguadé, M. *Genetics* **119**, 135-140 (1988).
20. Simmons, G. M., Kreitman, M., Quattlebaum, W. F. & Miyashita, N. *Evolution* **43**, 393-409 (1989).
21. Pasteur, N. & Georgiou, G. P. *J. Econ. Entomol.* **82**, 347-353 (1989).
22. Magnin, M., Marboutin, E. & Pasteur, N. *J. med. Ent.* **25**, 99-104 (1988).
23. Hemingway, J., Callaghan, A. & Amin, A. *Med. Vet. Ent.* **3**, 445-446 (1989).
24. Beyssat-Arnaouty, V., Mouchès, C., Georgiou, G. P. & Pasteur, N. *J. Am. Mosq. Control Assoc.* **5**, 196-200 (1989).
25. Urbanelli, S., Bullini, L. & Villani, F. *Bull. ent. Res.* **75**, 291-304 (1985).
26. Pasteur, N., Sinégre, G. & Gabinaud, A. *Biochem. Genet.* **19**, 499-508 (1981).
27. Villani, F. & Hemingway, J. *J. Pest. Biochem. Physiol.* **27**, 218-228 (1987).
28. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
29. Georgiou, G. P. & Pasteur, N. *J. Econ. Ent.* **73**, 489-492 (1980).
30. Georgiou, G. P., Metcalf, R. L. & Glidden, F. E. *Bull. WHO* **35**, 691-708 (1966).
31. Raymond, M. et al. *C. R. Acad. Sci. Paris* **300**, 509-512 (1985).
32. Beyssat-Arnaouty, V. thesis, Univ. Montpellier II (1989).
33. Southern, E. M. *J. molec. Biol.* **98**, 503 (1975).

ACKNOWLEDGEMENTS. We thank M. Marquine for technical assistance, G. P. Georgiou and I. Olivieri for providing strains, G.P.G. for allowing us to use his unpublished data, V. Beyssat, P. Borsia, V. Perrot, E. Poulin, L. Pouyaud, and M. A. Réglade for help in restriction mapping, and P. Jarne, C. Malcolm and Y. Michalakos for helpful discussions. A.C. was the recipient of a Royal Society/CNRS fellowship. This work was supported by PIREN/CNRS and MRT.

Action potentials must admit calcium to evoke transmitter release

Rosel M. Mulkey & Robert S. Zucker

Division of Neurobiology, Department of Molecular and Cell Biology, University of California, Berkeley, California 94720, USA

THERE are two hypotheses to explain how neurons release transmitter. The calcium hypothesis proposes that membrane depolarization is necessary only for opening calcium channels and increasing internal calcium concentration ($[Ca^{2+}]_i$) near membrane transmitter-release sites¹⁻³. These calcium ions trigger a transient release of neurotransmitter^{4,5}. The calcium-voltage hypothesis postulates that voltage induces a conformational change in a membrane protein rendering it sensitive to calcium such that, in the presence of high $[Ca^{2+}]_i$, depolarization directly triggers transmitter release⁶⁻⁹. Here we report that when calcium influx is blocked by cobalt or manganese ions in a calcium-free Ringer, as measured with Fura-2, and $[Ca^{2+}]_i$ is elevated by liberation from a caged calcium compound, transmitter release at the crayfish neuromuscular junction is unaffected by presynaptic action potentials. These results support the calcium hypothesis.

To distinguish between the hypotheses we had to independently increase $[Ca^{2+}]_i$ in terminals and depolarize the membrane without allowing calcium influx. New photolabile 'caged' calcium compounds decrease their affinity for calcium on exposure to ultraviolet light and permit the elevation of $[Ca^{2+}]_i$ in presynaptic terminals without depolarizing the membrane. A direct effect of voltage on transmitter release at crayfish neuromuscular junctions was inferred recently¹⁰ from postsynaptic responses to presynaptic action potentials in low-calcium solutions containing calcium-channel blockers, following elevation of presynaptic $[Ca^{2+}]_i$ with the photosensitive calcium chelator nitr-5 or the mitochondrial poison carbonyl cyanide *m*-chlorophenylhydrazine. Upon testing these low-calcium solutions used in ref. 10 by using the calcium indicator dye Fura-2, we found that calcium influx was not fully blocked, and transmitter release was clearly evoked by presynaptic action potentials at high stimulus frequencies. Therefore, the conclusion that voltage has a direct effect on transmitter release is invalid, because membrane depolarization activated calcium influx. We reexamined the calcium-voltage hypothesis at the crayfish neuromuscular junction using the photolabile calcium chelator DM-nitrophen and tested solutions containing calcium-channel blockers using the calcium indicator Fura-2 to be certain calcium influx was abolished during membrane depolarization.

The fluorescent dye Fura-2 was iontophoresed into the excitatory motor axon of the first walking leg of the crayfish, *Procambarus clarkii*, and intracellular $[Ca^{2+}]_i$ in presynaptic boutons was determined ratiometrically from signals obtained with 350 and 385 nm illumination. Excitatory postsynaptic potentials (e.p.s.ps) were recorded simultaneously in muscle fibres where the imaged terminals synapsed or in adjacent fibres. An extracellular electrode on an axonal branch recorded nerve field potentials to confirm action potential invasion of terminals. A nerve field potential was also sometimes observed in the postsynaptic recording. We used the same stimulation paradigm employed in ref. 10 (twin pulses, 20 ms apart delivered at 10 Hz) as well as high frequency stimulation at 100 Hz.

Figure 1a exemplifies experiments in which $[Ca^{2+}]_i$ rose by 87 ± 17 nM (mean $\pm$ s.d., 21 terminals in 10 preparations designated $N=21, 10$) during a 2-min twin-pulse stimulation in 'calcium-free Ringer's solution' containing 12.5 mM magnesium. Even though there was a small calcium influx at this frequency, transmitter release was undetectable, owing to the highly non-linear relationship between calcium and release¹¹ (Fig. 1b and

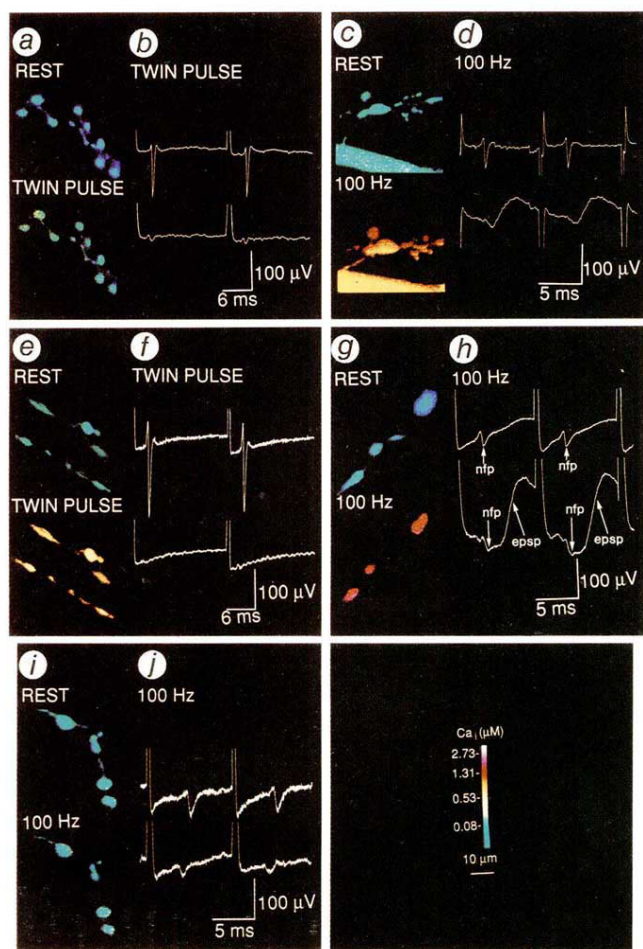


FIG. 1 Measurements of intracellular calcium and evoked transmitter release during presynaptic stimulation in 'calcium-free Ringer's' containing 12.5 mM Mg (a-d), 'manganese Ringer's' containing 0.2 mM Ca, 2 mM Mn, and 12.5 mM Mg (e-h), or 'cobalt Ringer's' containing 13.5 mM Co and 30 mM Mg (i, j). Each panel is from a different preparation. a, e, $[Ca^{2+}]_i$ measured at rest and during twin-pulse stimulation and c, g, i, 100 Hz stimulation. b, f, Averaged e.p.s.p.s (bottom trace) and nerve field potentials (n.f.p.s, top trace) recorded during twin-pulse stimulation and d, h, j, during 100 Hz stimulation. Neither calcium-free magnesium Ringer's nor manganese Ringer's with some calcium added prevented calcium influx and transmitter release; the cobalt Ringer's with no added calcium prevented both.

METHODS. The excitatory axon of the first walking leg of the crayfish (*Procambarus clarkii*), was injected with Fura-2 as described in ref. 13. Individual presynaptic terminals were imaged using a $\times 40$ Zeiss water immersion lens and Nikon Optiphot microscope equipped with a Dage silicon intensified target camera (Model 66). $[Ca^{2+}]_i$ was calculated from 'ratioed' fluorescence images with excitation at 350 and 385 nm (ref. 13). The excitatory axon was stimulated with a suction electrode placed on the nerve dissected from the meropodite. The e.p.s.p.s were recorded with an intracellular electrode containing 3 M KCl (10 M Ω) placed in a muscle fibre where imaged terminals synapsed. An extracellular electrode (25–50 μ m diameter), placed on an axonal branch on the surface of the opener muscle, recorded the nerve field-potential at its time of arrival in this particular branch. The e.p.s.p.s and motor nerve field-potentials were averaged on a digital oscilloscope (Nicolet 4094, Madison). Normal Ringer's contained (in mM) 195 NaCl, 13.5 CaCl₂, 5.4 KCl, 2.6 MgCl₂, and 10 sodium-HEPES at pH 7.4. Temperature was maintained at 18 °C. A gravity perfusion system with inflow above the opener muscle provided continuous perfusion with the appropriate saline solution. To ensure complete removal of calcium from the bath, the objective was lifted out of the bath and the muscle surface washed vigorously with a syringe having a 1-mm-diameter plastic tube attached to the tip. A perfusion time of 5–10 min was needed to eliminate calcium influx during a high-frequency train. We found that a manganese Ringer's with no added calcium was equivalent to the cobalt Ringer's in preventing calcium influx. When both calcium and manganese were present in the Ringer, a rise in $[Ca^{2+}]_i$ was measurable for only 20 s at 100 Hz. Manganese ions appear to pass through calcium channels in presynaptic terminals¹⁴ and then bind to Fura-2 and cause quenching of its fluorescence^{15–17}. From the intensity of Fura-2 fluorescence, we estimate its concentration at 100–150 μ M in boutons. Complete quenching suggests that about 100 μ M manganese accumulates during a 20 s, 100 Hz train. This may be compared to a rise in free $[Ca^{2+}]_i$ of over 5 μ M at 100 Hz in normal Ringer's¹³. About 99% of entering calcium is bound to cytoplasmic buffers¹⁸, so the influx of manganese through calcium channels is less than 20% that of calcium. We preferred the cobalt Ringer's because quenching of the Fura-2 signal took longer (40–60 s), suggesting that cobalt is less permeable than manganese through calcium channels.

ref. 10). At 100 Hz for 2 min, $[Ca^{2+}]_i$ increased from a resting value of 104 ± 45 nM to approximately 923 ± 313 nM ($N = 9, 5$) (Fig. 1c). This large increase in $[Ca^{2+}]_i$ facilitated transmitter release sufficiently that it was now easily detectable (Fig. 1d). When such experiments were done using a 'manganese Ringer's' containing 0.2 mM calcium, 2 mM manganese, and 12.5 mM magnesium, a large increase in $[Ca^{2+}]_i$ was seen at frequencies

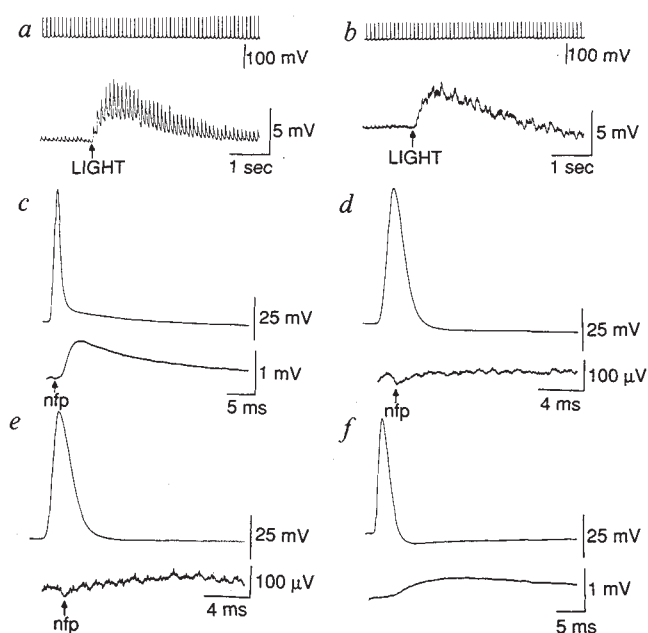


FIG. 2 Effect of action potentials on transmitter release in the absence of calcium entry. In each panel, the upper trace shows presynaptic action potentials stimulated at 10 Hz and the lower trace shows postsynaptic potentials. a, control experiment on a preparation injected presynaptically with DM-nitrophen and bathed in normal calcium solution. Photolysis with a 10-s exposure to UV light beginning at the arrow elevated presynaptic $[Ca^{2+}]_i$ and triggered transmitter release, seen as a depolarizing shift in the postsynaptic baseline. Spike-evoked e.p.s.p.s were also strongly facilitated. The gradual decline of the response probably reflects active extrusion of the photo-released calcium from the terminals and depletion of releasable transmitter quanta. The remaining five panels are from a different preparation showing that action potentials evoked no transmitter release in cobalt Ringer which effectively prevents calcium influx. b, Presynaptically injected nitrophen was photolysed to trigger transmitter release directly. Both before and during the light exposure, presynaptic spikes evoked no release in cobalt Ringer's. c–f, Averages of 50 responses to presynaptic stimulation in the experiment of panel b. c, Normal e.p.s.p.s evoked by action potentials in normal Ringer's. A small field potential reflecting invasion of the presynaptic terminals by action potentials preceded the e.p.s.p. d, E.p.s.p.s blocked by cobalt Ringer's but nerve terminal potential remaining, indicating that presynaptic terminals were still invaded. e, First 50 responses during photolysis of the nitrophen. The nerve terminal potential was still evident, and the increased noise reflected an acceleration in miniature e.p.s.p. frequency following elevation of $[Ca^{2+}]_i$. Nevertheless, action potentials evoked no transmitter release in the absence of calcium influx. When normal extracellular calcium was restored (f), action potentials again admitted calcium presynaptically, and e.p.s.p.s returned.

METHODS. Bevelled microelectrodes were filled with 60 mM DM-nitrophen, 18 or 36 mM CaCl₂, and 10 mM sulphorhodamine B (Molecular Probes) in 0.5 M potassium-HEPES, pH 7.3. The exciter axon was impaled near the Y-branch on the muscle surface and filled by pressure injection as described in ref. 19. By including 10 mM sulphorhodamine B (Molecular Probes) in the electrode, we could monitor filling the terminals without photolysing the DM-nitrophen, which reached a final concentration of about 3 mM. Light from a 100 W mercury epifluorescence illuminator was attenuated with a 25% transmission N.D. filter and focused onto terminals through a $\times 10$, 0.25 n.a. objective to photolysate about 20% of the nitrophen within 1 s, raising $[Ca^{2+}]_i$ from 0.1–0.5 μ M to several micromolar. The exact levels of $[Ca^{2+}]_i$ depend on concentrations of magnesium, ATP, DM-nitrophen, and native calcium buffer in nerve terminals¹⁹, and are presently undetermined. When nitrophen was injected without calcium loading in calcium-free media, photolysis evoked little or no transmitter release, indicating that release of magnesium from nitrophen does not trigger the responses. The injection electrode recorded the presynaptic action potential and a postsynaptic electrode was placed in a muscle fibre onto which filled terminals synapsed and from which a motor nerve field-potential could be recorded. Miniature e.p.s.p. frequency was estimated from the postsynaptic depolarization as described in ref. 5. A calcium-activated potassium current ($I_{K(Ca)}$) has been reported in crayfish presynaptic terminals²⁰. To prevent shunting of action potentials following $[Ca^{2+}]_i$ elevation by nitrophen photolysis, tetraethylammonium was added to cobalt Ringer's at 2 mM, a dose which is effective in blocking $I_{K(Ca)}$ (M. Wojtowicz, personal communication).

of 100 Hz, from 69 ± 25 nM at rest to 900 ± 606 nM ($N = 11, 5$) in 20 s (Fig. 1g), and synaptic potentials were also observed (Fig. 1h). Twin pulse stimulation for 90 s in manganese Ringer's produced a 189 ± 55 nM ($N = 8, 4$) increase in $[Ca^{2+}]_i$ (Fig. 1e) but, as in ref. 10, no detectable evoked transmitter release (Fig. 1f).

Since these calcium-free and manganese Ringer's permit calcium influx during membrane depolarization, this would cause evoked transmitter release which can be facilitated to detectable levels by elevating presynaptic $[Ca^{2+}]_i$, either through high frequency stimulation or the use of photolabile calcium chelators or mitochondrial poisons¹⁰. Therefore, such release does not support the calcium-voltage hypothesis. We tested this hypothesis by using the photolabile calcium chelator DM-nitrophen and solutions which more effectively prevent calcium entry. We chose DM-nitrophen, instead of nitr-5, because it binds calcium with a higher affinity before photolysis and a lower affinity after light exposure¹², providing bigger changes in $[Ca^{2+}]_i$. We found that a 'cobalt Ringer's' containing 13.5 mM cobalt, 30 mM magnesium, and no added calcium, permits only minimal calcium entry during action potentials. The $[Ca^{2+}]_i$ rose by only 64 ± 19 nM ($N = 11, 5$) during a 20-s 100-Hz stimulation, and no transmitter release was observed (Fig. 1i, j). The small rise in $[Ca^{2+}]_i$ and the presence of nerve field potentials indicate that action potentials were invading the nerve terminals but that calcium influx was insufficient to evoke transmitter release. A zero-calcium saline containing 2 mM EGTA was even more effective in reducing calcium influx but EGTA often led to presynaptic depolarization and failure of the action potential.

With a solution that effectively prevents calcium influx, it was possible to test the calcium-voltage hypothesis. In normal Ringer's, DM-nitrophen photolysis strongly increased transmitter release, expressed as the frequency of appearance of miniature e.p.s.ps, by 1,700–22,500% ($N = 12$); evoked e.p.s.ps were facilitated by 173–2,188%. These effects are much more powerful than those seen using nitr-5 (ref. 10).

Figure 2b–f illustrates one of six experiments to test for an effect of voltage on transmitter release without calcium influx. Initially, action potentials elicited large e.p.s.ps in normal Ringer's, in addition to a small nerve terminal potential recorded by the muscle electrode (Fig. 2c). After rinsing in cobalt Ringer's for 10 min, 10-Hz stimulation failed to evoke transmitter release, but the muscle electrode still recorded a nerve terminal potential (Fig. 2d). Photolysis of nitrophen triggered intense transmitter release, with miniature frequency rising from about 1 Hz to 4,500 Hz (800–10,700 Hz, $N = 6$) (Fig. 2b). Nevertheless, action potentials still failed to activate transmitter release (Fig. 2e). On return to normal Ringer's the e.p.s.p. reappeared in 3–4 min, smaller in amplitude, but broader than the control synaptic potential (Fig. 2f), due to the residual effects of TEA and incomplete exchange of solutions.

These and previous¹⁰ results lead us to the following conclusions about synaptic transmission at the crayfish neuromuscular junction: (1) saline solutions containing magnesium or manganese and low calcium permit calcium influx and presynaptic accumulation during action potentials and produce evoked release at high stimulus frequency; this transmitter release is potentiated by elevation of resting $[Ca^{2+}]_i$; (2) solutions containing cobalt (or manganese or EGTA) and no added calcium allow minimal calcium influx and as a result do not produce transmitter release to action potentials; (3) elevation of $[Ca^{2+}]_i$ in crayfish presynaptic boutons through photolysis of DM-nitrophen causes a large facilitation of evoked transmitter release and increases miniature e.p.s.p. frequency in a normal calcium medium; (4) in solutions which effectively block calcium influx during action potentials, elevation of $[Ca^{2+}]_i$ through photolytic release of 'caged' calcium strongly activates transmitter release, but the depolarization of an action potential produces no additional release; (5) the normal spike-evoked secretion of neurotransmitter is not affected directly by presynaptic voltage

but is triggered exclusively by an increase in $[Ca^{2+}]_i$ near transmitter release sites. □

Received 25 September; accepted 13 December 1990.

1. Zucker, R. S. & Landò, L. *Science* **231**, 574–579 (1986).
2. Zucker, R. S. & Haydon, P. G. *Nature* **335**, 360–362 (1988).
3. Zucker, R. S., Landò, L. & Fogelson, A. L. *J. Physiol., Paris* **81**, 237–245 (1986).
4. Katz, B. *The Release of Neural Transmitter Substances* (Thomas, Springfield, 1969).
5. Miledi, R. *Proc. R. Soc. B* **183**, 421–425 (1973).
6. Llinás, R., Steinberg, I. Z. & Walton, K. *Biophys. J.* **33**, 323–352 (1981).
7. Dudel, J., Parnas, I. & Parnas, H. *Pflügers Arch.* **399**, 1–10 (1983).
8. Parnas, I., Parnas, H. & Dudel, J. *Pflügers Arch.* **406**, 121–130 (1986).
9. Parnas, I., Parnas, H. & Dudel, J. *Pflügers Arch.* **406**, 131–137 (1986).
10. Hochner, B., Parnas, H. & Parnas, I. *Nature* **342**, 433–435 (1989).
11. Zucker, R. S. *A. Rev. Neurosci.* **12**, 13–31 (1989).
12. Kaplan, J. H. & Ellis-Davies, G. C. R. *Proc. natn. Acad. Sci. U.S.A.* **85**, 6571–6575 (1988).
13. Delaney, K. R., Zucker, R. S. & Tank, D. W. *J. Neurosci.* **9**, 3558–3567 (1989).
14. Narita, K., Kawasaki, F. & Kita, H. *Brain Res.* **510**, 289–295 (1990).
15. Hallam, T. J., Jacob, R. & Merritt, J. E. *Biochem. J.* **255**, 179–184 (1988).
16. Foder, B., Scharff, O. & Thastrup, O. *Cell Calcium* **10**, 477–490 (1989).
17. Grynkiewicz, G., Poenie, M. & Tsien, R. Y. *J. biol. Chem.* **260**, 3440–3450 (1985).
18. Fogelson, A. L. & Zucker, R. S. *Biophys. J.* **48**, 1003–1017 (1985).
19. Delaney, K. R. & Zucker, R. S. *J. Physiol.* **426**, 473–498 (1990).
20. Bittner, G. D., *J. Neurobiol.* **20**, 386–408 (1988).

ACKNOWLEDGEMENTS. We thank J. Kaplan and G. Ellis-Davies for DM-nitrophen, R. Y. Tsien for help in assembling the ratiometric imaging microscope, R. E. English for technical assistance, and K. Delaney for discussion. This research was supported by the NIH (R.S.Z. and R.M.).

Nature of the motor element in electrokinetic shape changes of cochlear outer hair cells

Peter Dallos, Burt N. Evans & Richard Hallworth

Auditory Physiology Laboratory (The Hugh Knowles Center) and Department of Neurobiology and Physiology, Northwestern University, Evanston, Illinois 60208, USA

IT is the prevailing notion that cochlear outer hair cells function as mechanical effectors as well as sensory receptors^{1–3}. Electrically induced changes in the shape of mammalian outer hair cells^{4,5}, studied *in vitro*, are commonly assumed to represent an aspect of their effector process that may occur *in vivo*. The nature of the motile process is obscure, even though none of the established cellular motors can be involved⁶. Although it is known that the motile response is under voltage control⁷, it is uncertain whether the stimulus is a drop in the voltage along the long axis of the cell or variation in the transmembrane potential. We have now performed experiments with cells partitioned in differing degrees between two chambers. Applied voltage stimulates the cell membrane segments in opposite polarity to an amount dependent on the partitioning. The findings show, in accordance with previous suggestions^{6,8}, that the driving stimulus is a local transmembrane voltage drop and that the cellular motor consists of many independent elements, distributed along the cell membrane and its associated cortical structures. We further show that the primary action of the motor elements is along the longitudinal dimension of the cell without necessarily involving changes in intracellular hydrostatic pressure. This establishes the outer hair cell motor as unique among mechanisms that control cell shape⁹.

Mammalian outer hair cells are slender cylindrical structures of fairly uniform diameter (~8–10 μ m) and their length ranges from ~20–30 μ m in the high-frequency cochlear base, to about 80–100 μ m in the low-frequency apex. These cells are distinguished by the presence of several layers of subsurface cisterns^{10,11} and an elaborate network of cortical cytoskeletal elements between the cisterns and the plasma membrane^{12–14}. Outer hair cells possess primarily efferent innervation with only a rudimentary afferent endowment^{15,16} that constitutes 5–10% of the cochlear outflow¹⁷. Yet, interference with outer hair cell function produces profound changes in afferent neural and behavioural responses^{18,19}, implying a mediating effect on

cochlear function. The electromotile mechanism⁴ is conceivably the vehicle for this mediation. Passive mechanisms of electrophoresis or electro-osmosis driven by longitudinal voltage drops along the cell have been proposed²⁰⁻²². An alternative suggestion is that transmembrane voltage drop drives the longitudinal motility through a change in circumference, coupled to longitudinal length change by the intermediate step of changing internal hydrostatic pressure^{6,8}.

In our experiments, isolated outer hair cells are inserted part-way into a pipette, designated the microchamber (Fig. 1), that is loosely patterned after the suction pipette used for examining retinal rods²³. The imposed voltage (V) drops on the series impedances of the included and extruded membrane segments, which form a voltage divider (Fig. 2a). Whereas current flow in the cytoplasm is unidirectional, it is inward through one membrane segment and outward through the other. Consequently, the voltage drops on the included and excluded membrane segments are in opposite polarity and have approximate magnitudes of pV and $(1-p)V$, where p is the fraction of the cell length extruded from the chamber. The key observation is shown in Fig. 3a. The two poles of the cell move in opposite phase; when the basal portion contracts, the apical portion elongates and vice versa. Displacements increase cumulatively from zero at the attachment of the cell to the microchamber, to maximum at the two cellular poles (Fig. 2b). These observations are independent of the waveform of the stimulus; here we demonstrate the response to a sinusoidal voltage command. Because the longitudinal intracellular voltage drop is unidirectional, whereas the included and extruded cell segments change their lengths in opposite phase, we conclude that the motor is driven by local membrane potential gradient.

A membrane-associated motor could work by one of two conceptually distinct mechanisms. The elementary motors could elongate or contract along the cell axis, thereby producing length changes directly. Alternatively the elementary motors could deform circumferentially⁶. This would produce increments (for radial contraction) or decrements (for radial expansion) in the cell's internal hydrostatic pressure. So that constant cell volume may be maintained, the cell would elongate or shorten. Support

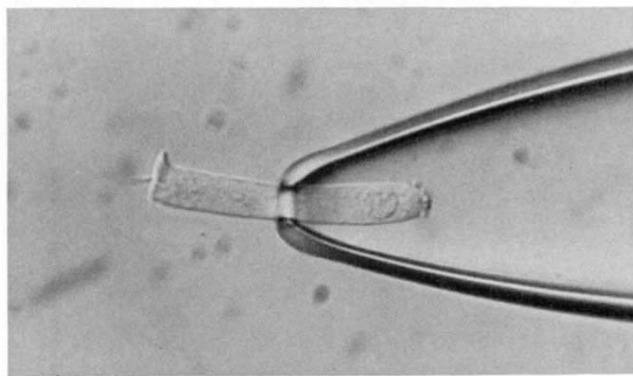


FIG. 1 Photographic view of the microchamber with outer hair cell partially inserted.

METHODS. Outer hair cells are isolated from guinea-pig cochleas by enzymatic digestion (medium is as in Leibovitz's L-15 but omitting everything except monovalent ion salts) and gentle trituration¹¹. The solution is supplemented by 15 mM HEPES buffer and 0.1% papain (Sigma) and adjusted to pH 6.9-7.0, osmolarity 310 mosmol. The microchamber is fabricated from borosilicate thin-wall glass tubing (2-mm outer diameter) pulled to a rapidly tapering tip of 10 μ m. The tip is then heat polished to a final inner diameter of 8 μ m. The final taper of the pipette is sufficiently rapid so that inserted cells are free of the chamber wall. The pipette is bent at 45° approximately 5 mm from the tip and coated with a silane compound to prevent adherence between chamber and cell. The microchamber is mounted in an experimental bath on the stage of a Zeiss inverted microscope.

for this possibility comes from observations showing that radial contraction-expansion cycles accompany longitudinal motility (ref. 8 and our own measurements). The microchamber method allows the direct testing of the two hypotheses. The number of motor elements in the extruded segment is proportional to p , whereas that in the included segment to $(1-p)$. Thus, if the longitudinal mechanism were correct, the length change of either extruded or included segments would be dependent on the product of the effective voltage and number of motor elements, $p(1-p)$, but would be in the opposite direction. This function is maximum when the cell is half inserted into the microchamber, $p=0.5$ (Fig. 3b). But if longitudinal motility relied on internal pressure change driven by circumferential motors, the motile response should be very small and the two ends ought to move in phase. This is expected because internal pressure changes, if they were produced in the two cell segments, would be equal and of opposite sign. Moreover, any net residual pressure change

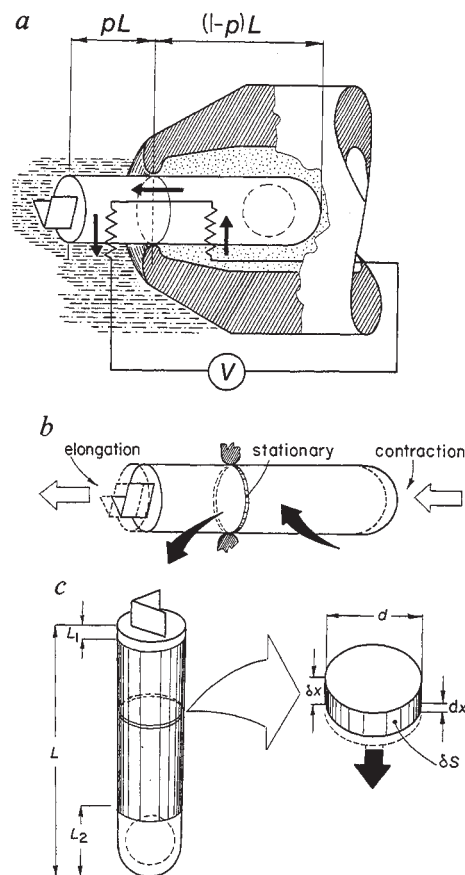
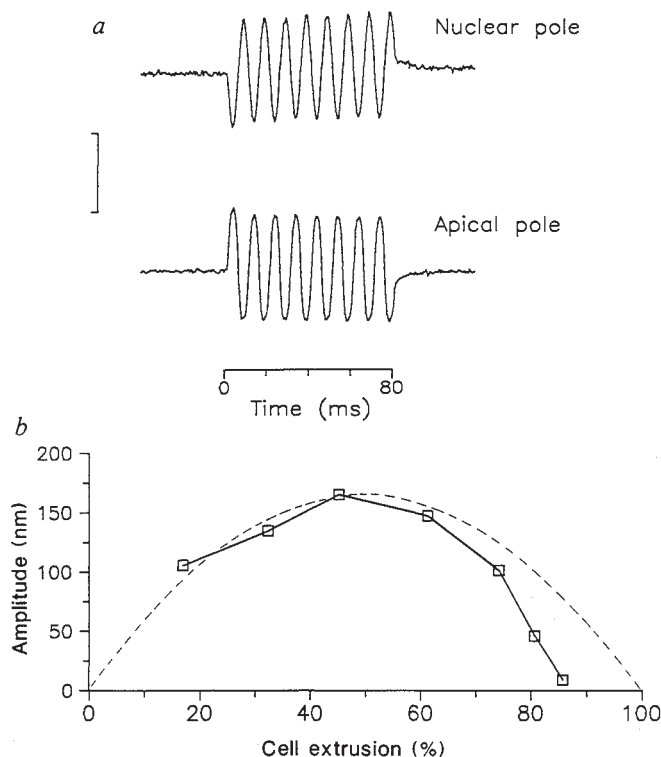


FIG. 2 a, Schematic view of the microchamber with a partially inserted cell. The voltage applied between the bath and the electrolyte filling the microchamber (the same solution in all these experiments: Leibovitz's L-15 (Gibco) buffered with 15 mM HEPES and adjusted to pH 7.35, 310 mOsmol) is V . b, Schematic of currents (black arrows) and resulting motions of the cell's two poles. When one end shortens, the other extends. This appears to the observer as if both ends moved in the same direction, that is, as if the cell moved in and out of the microchamber. But by closely observing cytoplasmic landmarks we satisfied ourselves that there is no motion at the narrow ring where the cell is attached to the microchamber. Noting the movement of various cellular features, it is clear that movement amplitude cumulatively increases away from the attachment site. c, Elementary circumferential ring, assumed to be the 'motor unit', of surface area δS and length δx . A voltage drop δV across the surface is presumed to result in longitudinal contraction (for depolarization) or extension (for hyperpolarization) of magnitude dx . Total motion is assumed to be the sum of elementary motions: $x = \sum dx$ ^{6,25}. The motor elements are taken to be coincident with the cell cortex that relates to subsurface cisterns and cortical lattice ($L - L_1 - L_2$). The region L_1 corresponds to the cuticular plate, whereas L_2 to the nuclear and synaptic region.



would be uniform for the whole cell, predicting that length changes for each segment would be in phase for any p . In Fig. 3b we plot the magnitude of motile response measured for the basal (included) pole of a cell, in response to a standard stimulus, with variable extrusion, p . It is apparent that maximal response occurs at some intermediate value of p . A similar result is obtained for the cell's other pole. We have seen no exception to this behaviour in over 20 cells studied. The critical observation, for the present purpose, is that maximal motility occurs somewhere towards the midpoint, $p \approx 0.5$, where the two cell segments, included and excluded, move in opposite phase. It has been shown that outer hair cells are normally under a positive intracellular resting pressure⁵ and, moreover, that this pressure is necessary for normal electromotility⁶. We do not dispute these findings, rather we state that change about this resting pressure is not an intermediate step of electromotility. It is further emphasized that the quantitative dependence of the motile response on the degree of cell extrusion from the microchamber is the strongest indication yet obtained of the distributed nature of the elementary motors. Because antiphasic responses between included and extruded segments can be elicited independent of the point of partitioning anywhere within the cell's supranuclear region, the motors must function independently and the length change of a cell or cell segment is simply the sum of the length changes of the individual motor elements. The findings rule out mechanisms of motility that depend on longitudinal voltage gradients such as electro-osmosis²¹ and certain forms

of electrophoresis²². The experiments described here clearly demonstrate that outer hair cell electromotility is driven by a large number of small, independent, voltage-sensing⁷ elementary motors, closely associated with the plasma membrane, whose primary mode of action is longitudinal deformation, as shown schematically in Fig. 2c. The anatomical basis for the mechanism remains unknown. It is conceivable that transmembrane voltage gradients produce conformational changes in protein structures associated with the cell's cortex that directly change the longitudinal length of the elementary motor units. Longitudinally arranged crossbridges²⁴ linking helical circumferential filaments ('springs'¹³) of the cortical lattice¹⁴ are ideally situated to convey or generate forces in the axial direction, as has been suggested²⁴. This possibility is consistent with the distributed nature of the motor elements and the summation of their contributions^{6,11,25} and their association with the cell cortex in contrast to some intracellular structure^{6,25}. Because small changes in diameter are also seen during electromotility, it is conceivable that the motors and their attachments are so arranged that they yield vectorial force components in both axial and radial directions. That the elementary motors primarily act in the longitudinal direction implies that they are in effect connected 'in series'. In the condition of these experiments, in the absence of an external mechanical load, the cell behaves as a displacement generator and it is conceivable that it may function in a similar manner *in vivo*. □

of electrophoresis²².

The experiments described here clearly demonstrate that outer hair cell electromotility is driven by a large number of small, independent, voltage-sensing⁷ elementary motors, closely associated with the plasma membrane, whose primary mode of action is longitudinal deformation, as shown schematically in Fig. 2c. The anatomical basis for the mechanism remains unknown. It is conceivable that transmembrane voltage gradients produce conformational changes in protein structures associated with the cell's cortex that directly change the longitudinal length of the elementary motor units. Longitudinally arranged crossbridges²⁴ linking helical circumferential filaments ('springs'¹³) of the cortical lattice¹⁴ are ideally situated to convey or generate forces in the axial direction, as has been suggested²⁴. This possibility is consistent with the distributed nature of the motor elements and the summation of their contributions^{6,11,25} and their association with the cell cortex in contrast to some intracellular structure^{6,25}. Because small changes in diameter are also seen during electromotility, it is conceivable that the motors and their attachments are so arranged that they yield vectorial force components in both axial and radial directions. That the elementary motors primarily act in the longitudinal direction implies that they are in effect connected 'in series'. In the condition of these experiments, in the absence of an external mechanical load, the cell behaves as a displacement generator and it is conceivable that it may function in a similar manner *in vivo*. □

Received 29 October; accepted 11 December 1990.

1. Davis, H. *Hearing Res.* **9**, 79–90 (1983).
2. Dallos, P. in *Contemporary Sensory Neurobiology* (eds Correia M. J. & Perachio A. A.) 207–230 (Alan R. Liss, New York, 1985).
3. Kim, D. O. *Hearing Res.* **22**, 105–114 (1986).
4. Brownell, W. E. in *Mechanisms of Hearing* (eds Webster W. R. & Aitkin L. M.) 5–10 (Monash Univ. Press, Clayton, Australia, 1983).
5. Brownell, W. E., Bader, C. R., Bertrand, D. & de Ribaupierre, Y. *Science* **227**, 194–196 (1985).
6. Holley, M. C. & Ashmore, J. F. *Proc. R. Soc. B* **232**, 413–429 (1988).
7. Santos-Sacchi, J. & Dilger, J. P. *Hearing Res.* **35**, 143–150 (1988).
8. Ashmore, J. F. *J. Physiol., Lond.* **388**, 323–347 (1987).
9. Lackie, J. M. *Cell Movement and Cell Behaviour* (Allen & Unwin, London, 1986).
10. Saito, K. *Cell Tiss. Res.* **229**, 467–481 (1983).
11. Evans, B. N. *Hearing Res.* **45**, 265–282 (1990).
12. Flock, A., Flock, B. & Ulfendahl, M. *Arch. Otorhinolaryngol.* **243**, 83–90 (1986).
13. Holley, M. C. & Ashmore, J. F. *Nature* **335**, 635–637 (1988).
14. Bannister, L. H., Dodson, H. C., Astbury, A. F. & Douek, E. E. *Prog. Brain Res.* **74**, 213–219 (1988).

15. Smith, C. A. & Dempsey, E. W. *Am. J. Anat.* **100**, 337–367 (1957).
16. Guinan, J. J., Warr, W. B. & Norris, B. E. *J. comp. Neurol.* **221**, 358–370 (1983).
17. Spoendlin, H. *Acta Otolaryngol. (Stockholm)* **67**, 239–254 (1969).
18. Dallos, P. & Harris, D. M. *J. Neurophysiol.* **41**, 365–383 (1978).
19. Dallos, P. in *Auditory Function* (eds Edelman G. M., Gall W. E. & Cowan W. M.) 153–188 (J. Wiley, New York, 1988).
20. Brownell, W. E. & Kachar, B. in *Peripheral Auditory Mechanisms* (eds Allen J. B., Hall J. L., Hubbard A., Neely S. T. & Tubis A.) 369–376 (Springer-Verlag, New York, 1986).
21. Brownell, W. E. *Ear and Hearing* **11**, 82–92 (1990).
22. Jen, D. H. & Steele, C. R. *J. acoust. Soc. Amer.* **82**, 1667–1678 (1987).
23. Baylor, D. A., Lamb, T. D. & Yau, K.-W. *J. Physiol., Lond.* **288**, 589–611 (1979).
24. Holley, M. C. & Ashmore, J. F. *J. Cell Sci.* **96**, 283–291 (1990).
25. Evans, B. N. *thesis*, Univ. Texas Health Sciences Center, Houston (1988).
26. Clark, B. A., Hallworth, R. & Evans, B. N. *Pflügers Arch.* **415**, 490–493 (1990).

ACKNOWLEDGEMENTS. We thank M. A. Cheatham, B. Clark, S. Echter, M. Killian and J. Siegel for their comments on the manuscript. This work is supported by the NIH and the American Hearing Research Foundation.

Tyrosine phosphorylation and tyrosine kinase activity of the *trk* proto-oncogene product induced by NGF

David R. Kaplan*, Dionisio Martin-Zanca†‡
& Luis F. Parada†

* Eukaryotic Signal Transduction Group

† Molecular Embryology Group, ABL-Basic Research Program, NCI-Frederick Cancer Research and Development Center, PO Box B, Frederick, Maryland 21701, USA

‡ Present address: Departamento de Microbiología, Universidad de Salamanca, 37008 Salamanca, Spain

NERVE growth factor (NGF) is a neurotrophic factor responsible for the differentiation and survival of sympathetic and sensory neurons as well as selective populations of cholinergic neurons^{1,2}. NGF binds to specific cell-surface receptors but the mechanism for transduction of the neurotrophic signal is unknown. Several experiments using the NGF-responsive pheochromocytoma cell line, PC12, have implicated tyrosine phosphorylation in NGF-mediated responses, although no NGF-specific tyrosine kinases have been identified³. Here we show that NGF induces tyrosine phosphorylation and tyrosine kinase activity of the *trk* proto-oncogene product, a tyrosine kinase receptor whose expression is restricted *in vivo* to neurons of the sensory spinal and cranial ganglia of neural crest origin⁴. Tyrosine phosphorylation of *trk* by NGF is rapid, specific and occurs with picomolar quantities of factor, indicating that the response is mediated by physiological amounts of NGF. Activation of the *trk* tyrosine kinase receptor provides a possible mechanism for signal transduction by NGF.

In the presence of NGF, PC12 cells differentiate into sympathetic neurons, extend neurites, increase neurotransmitter

synthesis and become electrically excitable⁵⁻⁷. Several unidentified proteins are phosphorylated on tyrosine in these cells and the introduction of the tyrosine kinase *v-src* into PC12 cells stimulates an outgrowth of neurites similar to that resulting from NGF treatment^{3,8}. Therefore, events mediated by tyrosine phosphorylation could be involved in the differentiation of PC12 cells.

The *trk* gene is a member of a gene family of tyrosine kinase (TK) receptors that includes the related gene *trkb*^{4,9,10}. To determine if *trk* or *trkb* were transcribed in PC12 cells, the expression of *trk* transcripts was assayed by northern-blot transfer analysis with a full-length *trk* complementary DNA probe or with an extracellular domain, *trkb*-specific probe⁹. PC12 cells expressed a 3.2-kilobase *trk* transcript, but *trkb* transcripts were not detected (Fig. 1a). The level of *trk* transcripts was not affected by the addition of NGF. To determine whether additional *trk*-related genes were transcribed in PC12 cells, messenger RNA was hybridized at low stringency with the highly conserved *trk* TK domain. No additional *trk*-hybridizing transcripts were detected (not shown).

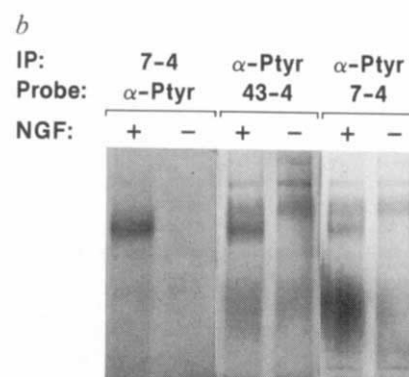
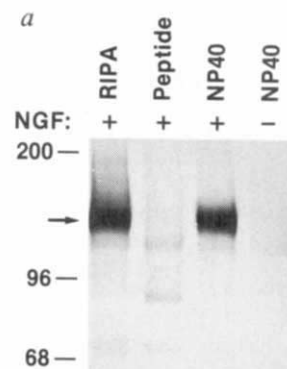


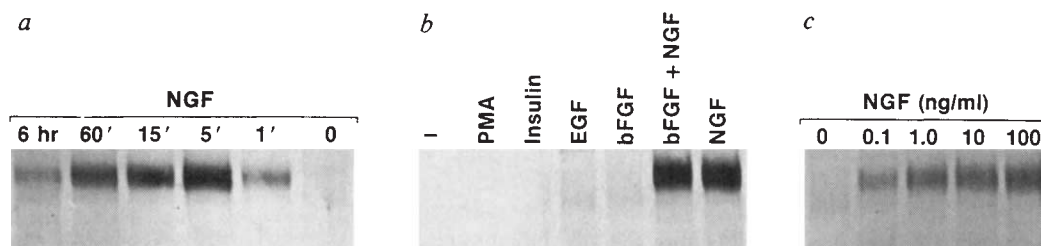
FIG. 1 a, Northern-blot transfer analysis of *trk* and *trkb* transcripts in NGF treated (+) or untreated (–) PC12 cells. b, Ptyr: phosphotyrosine-containing proteins in PC12 cells treated for 5 min or 5 days with NGF or in untreated cells (0). WG, tyrosine-phosphorylated glycoproteins in cells treated with NGF for 5 min or 5 days. *M_r* markers (in K) are indicated.

METHODS. a, RNA preparation and northern blot transfer analysis was performed as described previously⁴. Cells were treated with 50 ng per ml NGF (+) (Boehringer–Mannheim) and were collected 48 h later after differentiation had occurred. Total RNA (20 µg) was loaded per lane and the filter was hybridized with a probe specific for *trk*⁴ or *trkb*⁹. b, Ptyr: PC12 cells were treated with 50 ng per ml NGF for 5 min or 5 days. Cells were washed, lysed and the lysates (50 µg protein per lane) analysed by immunoblotting with the phosphotyrosine (Ptyr) monoclonal antibody 4G10 (gift from D. Morrison, B. Drucker and T. M. Roberts) as previously described^{17,18}. For WG, lysates (prepared from 2×10^7 cells) were incubated with 100 µl of a 50% solution of wheat-germ (WG) lectin agarose (Pharmacia) for 2 h at 4 °C. Precipitates were washed and analysed by western blot analysis with anti-Ptyr as previously described^{17,18}. Electrophoresis was performed on 7.5% SDS-polyacrylamide gels.

FIG. 2 NGF induces tyrosine phosphorylation of Trk. a, Trk was immunoprecipitated with rabbit peptide antibody 43-4¹¹ from lysates of NGF-treated (+) or untreated (–) PC12 cells and the immunoprecipitates were probed with anti-Ptyr. Cells were lysed in buffer containing 1% NP40¹⁸ or RIPA buffer¹⁹. Immunoprecipitations were performed in the presence or absence of competing peptide (10 µg ml^{–1}). Similar results were obtained with other sources of NGF (Collaborative Research, Upstate Biotechnology Inc.). b, NGF-induced tyrosine phosphorylation of Trk in anti-Trk or anti-Ptyr immunoprecipitates. Trk immunoprecipitates were prepared with rabbit antibody 7-4¹¹ before anti-Ptyr western blot analysis or with Ptyr antibody before western blot analysis with Trk antibodies 43-4 or 7-4.

METHODS. PC12 cells were treated for 5 min at 37 °C with 50 ng per ml NGF. Cells (2×10^7) were lysed in buffer containing 1% Nonidet P-40 (NP40) (except as noted) and were immunoprecipitated with anti-Trk or anti-Ptyr for 2–3 h at 4 °C as described previously¹⁸. Immunoprecipitates were collected with protein-A-Sepharose, subjected to 7.5% SDS-PAGE and analysed by immunoblotting as described¹⁸. *M_r* markers (in K) are indicated.

FIG. 3 Time course, growth factor specificity and dose-response of Trk tyrosine phosphorylation in PC12 cells. **a**, Time course of Trk tyrosine phosphorylation. Cells (2×10^7) were treated with 50 ng per ml NGF at 37 °C. **b**, Effects of growth and differentiation factors on Trk tyrosine phosphorylation. Cells were treated with 100 ng per ml NGF, 100 ng per ml basic fibroblast growth factor (bFGF) (Boehringer-Mannheim), 100 ng per ml epidermal growth factor (EGF) (Upstate Biotechnology, Inc.), 100 nM insulin (Sigma), or 1 μ g per ml phorbol 12-myristate 13-acetate (PMA) (Sigma) for 5 min at 37 °C. **c**, Dose-response of Trk tyrosine phosphorylation. Cells were treated



for 30 min at 37 °C with increasing concentrations of NGF. Western blot analysis with Ptkr antibodies of Trk immunoprecipitates prepared with anti-Trk antibody 43-4 are shown.

Phosphotyrosine (Ptyr) western-blot transfers of cell lysates prepared from PC12 cells treated with NGF for 5 min demonstrated the tyrosine phosphorylation of a broad band at relative molecular mass 130,000–140,000 (M_r 130–140K) (Fig. 1b). Proteins of 42, 45, 50, 60 and 150K were also observed to be phosphorylated on tyrosine in response to NGF treatment of PC12 cells. A band of 130–140K was observed when the cell lysates were precipitated with wheat-germ lectin before analysis by anti-Ptyr immunoblotting, indicating that NGF induced the tyrosine phosphorylation of a glycosylated protein (Fig. 1b).

To determine whether Trk, a 140K glycoprotein¹¹, was phosphorylated on tyrosine in response to NGF, Trk protein was immunoprecipitated with an antibody generated against a C-terminal Trk peptide (43-4)¹¹ and the immunoprecipitates were probed with anti-Ptyr antibodies. Trk was phosphorylated on tyrosine after addition of NGF to the cells (Fig. 2a). Immunoprecipitation of Trk was specifically blocked by a Trk-derived peptide and was not reduced by immunoprecipitation in stringent buffers such as RIPA (Fig. 2a). Tyrosine phosphorylation of Trk was also observed when the immunoprecipitations were performed using a second anti-Trk antibody generated against the p70^{trk} oncogene product (7-4)¹¹ and when anti-Ptyr immunoprecipitates prepared from NGF-treated cells were probed with either of two anti-Trk antibodies (Fig. 2b). The same amount of Trk protein was recovered in each immunoprecipitation, as measured by [³⁵S] methionine labelling or Trk (not shown). To confirm that Trk was phosphorylated on tyrosine residues in response to NGF, PC12 cells were labelled with ³²P-orthophosphate before NGF addition and immunoprecipitation with anti-Trk antibodies. Phosphoamino-acid analysis demonstrated that Trk was phosphorylated on tyrosine in response to NGF (not shown).

Trk tyrosine phosphorylation occurred within 1 min of NGF

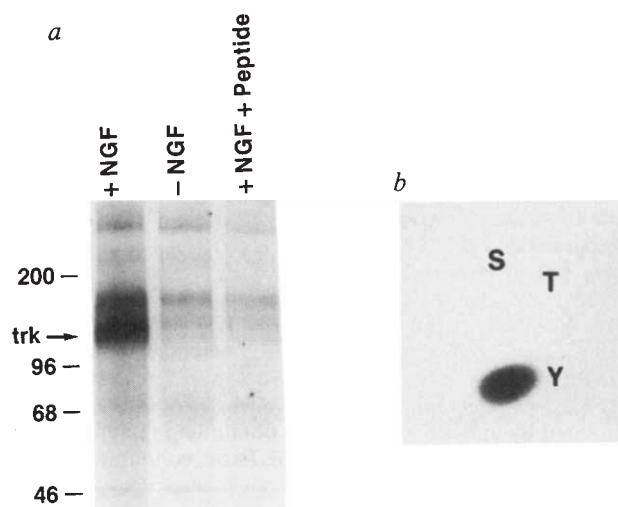
treatment of cells, reached maximum levels after 5 min and declined thereafter (Fig. 3a). Residual phosphorylation was detected after two days of treatment with NGF when the cell population was fully differentiated (not shown). Trk tyrosine phosphorylation was also specific to NGF. Other peptide growth factors that elicit tyrosine phosphorylation in PC12 cells were tested in our assay^{3,12}. Epidermal growth factor (EGF), basic fibroblast growth factor (bFGF), insulin and the phorbol ester phorbol 12-myristate 13-acetate (PMA) failed to induce Trk tyrosine phosphorylation (Fig. 3b). No additional tyrosine phosphorylation of Trk was seen in cells treated with both bFGF and NGF (Fig. 3b). These agents produce similar patterns of early responses in PC12 cells, including transcriptional activation of *c-fos* and *c-myc*¹³. But of these factors, only NGF and bFGF stimulate neurite outgrowth.

To determine the minimal concentration of NGF required for Trk tyrosine phosphorylation, a dose-response experiment was performed. Tyrosine phosphorylation was half-maximal at 0.1 ng per ml NGF (50 pM) (Fig. 3c), indicating that Trk phosphorylation occurs at physiologically relevant concentrations of NGF¹⁴.

NGF also induces the tyrosine kinase activity of Trk. Trk was phosphorylated *in vitro* in immunoprecipitates prepared from PC12 cells treated for 5 min with NGF, presumably owing to catalytic autophosphorylation (Fig. 4). No Trk phosphorylation was detected in immunoprecipitates from untreated cells or when competing Trk-derived peptide was included during immunoprecipitation.

Our results show that the protein tyrosine kinase receptor Trk is a substrate for an NGF-induced tyrosine kinase activity in PC12 cells. The rapid and specific response suggests that Trk is closely associated with the receptor for NGF. Alternatively, the Trk protein may directly bind NGF. The kinase activity of Trk

FIG. 4 Activation of Trk kinase activity by NGF. **a**, Immune-complex kinase assay following immunoprecipitation with anti-Trk antibodies prepared from NGF-treated (+) or untreated (–) PC12 cells. Immunoprecipitations were performed in the presence or absence of competing peptide (10 μ g ml^{–1}). **b**, Phosphoamino-acid analysis autoradiogram of Trk phosphorylated *in vitro* in **a**. The positions of serine (S), threonine (T) and tyrosine (Y) are indicated. **METHODS**. PC12 cells (8×10^7) were lysed and precipitated with wheat-germ lectin-agarose (300 μ l 50% solution) as for Fig. 1. The precipitates were washed three times in 1% NP40-containing lysis buffer¹⁸ and the glycoproteins were eluted in 400 μ l 0.5 M *N*-acetylglucosamine in lysis buffer containing 1 mM orthovanadate for 1 h at 4 °C. The eluate was diluted 10-fold with lysis buffer and immunoprecipitated with anti-Trk antibody as described in Fig. 2 legend. The immunoprecipitates were incubated with 20 μ Ci [γ -³²P] ATP, 10 mM MnCl₂, 20 mM Tris, pH 7.4, for 20 min at 25 °C. Phosphorylated proteins were analysed by 7.5% SDS-PAGE as described¹⁸. M_r markers (in K) are indicated. For **b**, the phosphorylated Trk band was eluted from the gel and phosphoamino-acid analysis performed as previously described²¹.



may also be activated by phosphorylation of Trk by another cellular tyrosine kinase.

Trk may be responsible for the increases in phosphotyrosine in NGF-treated PC12 cells. Several proteins phosphorylated on tyrosine in these cells are candidate substrates for the Trk-induced kinase activity, including microtubule associated protein-2 (MAP-2) kinase and phospholipase C γ ¹⁵ (M. Vetter, D.M.-Z., L.F.P., J. M. Bishop and D.R.K., manuscript submitted). Examination of Trk immunoprecipitates for co-immunoprecipitating substrates should facilitate the identification of other targets of tyrosine kinase activities in PC12 cells.

The expression *in vivo* of *trk* is consistent with its proposed role in NGF-mediated events. The *trk* mRNA is restricted to the dorsal root ganglia and cranial sensory ganglia of neural crest origin, tissues comprised primarily of neurons that are trophic substrates for NGF⁴. These neurons rely on NGF for differentiation and maintenance of the neuronal phenotype¹⁶. Sympathetic neurons and certain neurons in the central nervous system, which may not express *trk*, also respond to NGF^{1,2}. But other *trk* family members, for example *trkb*, are expressed in the central nervous system^{9,10} and may function as NGF-responsive tyrosine kinases in these tissues. Other tyrosine kinases expressed in neurons, including p60^{c-src} and p62^{c-yes}, may be involved in NGF responses, although the phosphorylation state and kinase activity of these proteins does not seem to be altered in response to addition of NGF (M. Vetter, D.K. and M. J.

Bishop, unpublished results). Trk is a likely candidate for a transducer of NGF signals. These results implicate tyrosine kinase receptors as possible mediators of neurotrophic signals in vertebrates. □

Received 18 January; accepted 12 February 1991.

1. Levi-Montalcini, R. *Science* **237**, 1154–1161 (1987).
2. Gnahn, H., Heft, F., Heumann, R., Schwab, M. E. & Thoenen, H. *Dev. Brain Res.* **9**, 45–52 (1983).
3. Maher, P. A. *Proc. natn. Acad. Sci. U.S.A.* **85**, 6788–6791 (1988).
4. Martin-Zanca, D., Barbacid, M. & Parada, L. F. *Genes Dev.* **4**, 683–694 (1990).
5. Greene, L. A. & Tischler, A. S. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2424–2428 (1976).
6. Schubert, D., Heinemann, S. & Kidokoro, Y. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2579–2583 (1977).
7. Dichter, M. A., Tischler, A. S. & Greene, L. A. *Nature* **268**, 501–504 (1977).
8. Alema, S., Casalboni, P., Agostini, E. & Tatro, F. *Nature* **316**, 557–559 (1985).
9. Klein, R., Martin-Zanca, D., Barbacid, M. & Parada, L. F. *Development* **4**, 845–850 (1990).
10. Klein, R., Parada, L. F., Coulier, F. & Barbacid, M. *EMBO J.* **8**, 3701–3709 (1989).
11. Martin-Zanca, D., Oskam, R., Mitra, G., Copeland, T. & Barbacid, M. *Molec. cell. Biol.* **9**, 24–33 (1989).
12. Maher, P. A. *J. Neurosci. Res.* **24**, 29–37 (1989).
13. Greenberg, M. E., Hermanowski, A. L. & Ziff, E. B. *Molec. cell. Biol.* **6**, 1050–1057 (1986).
14. Greene, L. A. & Rein, G. *Nature* **268**, 349–351 (1977).
15. Kim, U.-H. et al. *J. Biol. Chem.* (in the press).
16. Thoenen, H. & Barde, Y. A. *Physiol. Rev.* **60**, 1284–1335 (1980).
17. Morrison, D. K., Browning, P. J., White, M. F. & Roberts, T. M. *Molec. cell. Biol.* **8**, 176–185 (1988).
18. Kaplan, D. R., Morrison, D. K., Wong, G., McCormick, F. & Williams, L. T. *Cell* **61**, 125–133 (1990).
19. Morrison, D. K. et al. *Cell* **58**, 649–657 (1989).
20. Chao, M. V. in *Handbook of Experimental Pharmacology* Vol. 95/II (eds Sporn, M. B. & Roberts, A. B.) 135–165 (Springer-Verlag, 1990).
21. Sefton, B. M., Hunter, T., Ball, E. H. & Singer, S. J. *Cell* **24**, 165–174 (1981).

ACKNOWLEDGEMENTS. We thank D. Morrison for phosphoamino acid analysis, H. Varmus for support and advice, and M. Chao, B. Hempstead and M. Vetter for helpful discussions. We also thank N. Copeland, S. Hughes, M. Chao and B. Hempstead for critically reading the manuscript, J. Maragos and S. Rabin for technical help, and R. Handley for assistance in preparing the manuscript. This research was supported by the National Cancer Institute, DHHS, under contract.

A cellular protein that competes with SV40 T antigen for binding to the retinoblastoma gene product

Shi Huang, Wen-Hwa Lee & Eva Y.-H. P. Lee*

Department of Pathology, School of Medicine, University of California, San Diego, La Jolla, California 92093-0612, USA

TUMOUR-suppressor genes, such as the human retinoblastoma susceptibility gene (*Rb*), are widely recognized as being vital in the control of cell growth and tumour formation¹. This role is indicated, in part, by the suppression of tumorigenicity of human tumour cells after retrovirus-mediated *Rb* replacement^{2–4}. How *Rb* acts to bring about this suppression is not clear⁵ but one clue is that the *Rb* protein forms complexes with the transforming oncoproteins of several DNA tumour viruses^{6–8}, and that two regions of *Rb* essential for such binding frequently contain mutations in tumour cells^{9,10}. These observations suggest that endogenous cellular proteins might exist that bind to the same regions of *Rb* and thereby mediate its function. We report here the identification of one such human cellular *Rb*-associated protein of relative molecular mass 46,000 (46K) (RbAP46). Two lines of evidence support the notion that RbAP46 and simian virus 40 T antigen have homologous *Rb*-binding properties: first, several mutated *Rb* proteins that failed to bind to T also did not associate with RbAP46; and second, both T antigen and T peptide (amino acids 101–118) were able to compete with RbAP46 for binding to *Rb*. The apparent targeting of the RbAP46–*Rb* interaction by oncoproteins of DNA tumour viruses strongly suggests that formation of this complex is functionally important.

Previous attempts to identify *Rb*-associated proteins¹¹ may have been hampered by insufficient quantities of such stable complexes in cells. One approach to overcome this difficulty is to add excess free *Rb* protein *in vitro*, thereby driving complex formation. Detectable amounts of RbAPs could then be coprecipitated by anti-*Rb* antibody. To obtain large quantities of *Rb* protein, we constructed plasmid pETRbc, containing the entire

coding region of *Rb*, for high-level expression in *Escherichia coli*¹² (Fig. 1a). The most stably expressed protein was a 56K polypeptide (p56-Rb), which was then purified (Fig. 1b) and characterized immunologically (Fig. 1c). Only antibodies recognizing the C-terminal half of *Rb* could precipitate this protein, whereas two different antibodies recognizing the N-terminal regions failed to precipitate it. The C-terminal 56K contains both regions essential for T binding^{9,10}. When purified T antigen was added to bacterial lysates, this 56K polypeptide was coprecipitated by anti-T antibody (data not shown). Conversely, mixing increasing amounts of purified 56K protein with radiolabelled COS cell lysates, in which T is expressed in great excess over endogenous *Rb*⁷, resulted in quantitative coprecipitation of T by anti-*Rb* antibody over a 10-fold range (data not shown). Similar results were obtained by mixing COS cell extracts with lysates of bacteria expressing *Rb* protein (Fig. 3, lanes 1 and 2). These studies indicated that bacterially expressed p56-Rb retained the T-binding properties of full-length *Rb*.

We next tested whether a cellular protein exists in human HeLa cell extracts that binds to p56-Rb in a similar way. As shown in Fig. 2a, a protein of 46K (RbAP46) was specifically detected only when p56-Rb, either in bacterial lysates (Fig. 2a, lane 2) or in a purified form (Fig. 2a, lanes 4–7), was mixed with radiolabelled HeLa cell lysates. By mixing increasing amounts of purified p56-Rb with constant amounts of HeLa cell lysates, the 46K protein was immunoprecipitated in increasing amounts (Fig. 2a, lanes 3–7), suggesting that it formed a complex with the input 56K protein. To exclude the possibility of cross-reaction of the 46K protein with the rabbit anti-*Rb* 0.495, another anti-*Rb* antibody, 0.47, was used¹⁰, and again the 46K protein was coprecipitated (Fig. 2b, lane 11). But no 46K protein was detected if nonimmune serum was used (Fig. 2b, lane 9). These results implied that this 46K protein was directly associated with p56-Rb. Using a full-length *Rb* protein expressed in insect cells¹³, the same 46K protein was again coprecipitated (Fig. 2c, lane 15), although some additional protein bands were also seen. As RbAP46 was the major protein detected by using either a full-length *Rb* or p56-Rb, we concluded that p56-Rb contains all the necessary sequences for binding to the 46K protein.

The specificity of the interaction was further characterized by defining the regions of *Rb* involved in binding to RbAP46. Five

* To whom correspondence should be addressed.

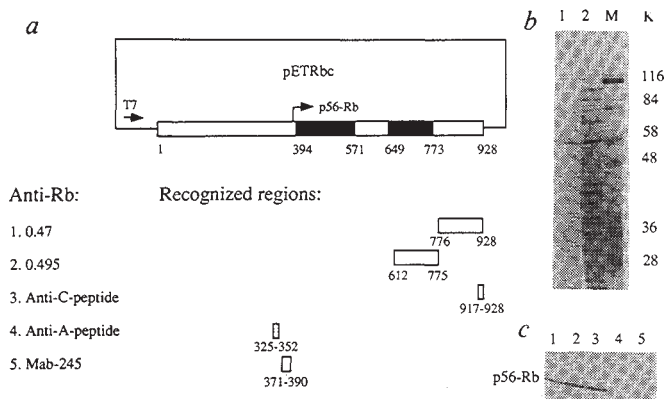


FIG. 1 Expression of Rb polypeptides in *E. coli*. **a**, Maps of plasmid pETRbc and regions of Rb protein recognized by various Rb antibodies. T-binding regions of Rb are represented as filled squares. **b**, Purification of p56-Rb. Bacterial lysate containing p56-Rb (lane 2) and one of the peak fractions collected from CM-Sepharose chromatography (lane 1) were analysed by 10% SDS-PAGE and stained with Coomassie blue. Lane M shows migration of markers, molecular weights indicated (in thousands) to the right. **c**, Immunological characterization of p56-Rb. Partially purified p56-Rb was immunoprecipitated with various anti-Rb antibodies (lane 1, 0.47; lane 2, 0.495; lane 3, anti-Rb c-peptide (amino acids 917-928); lane 4, anti-Rb A-peptide (amino acids 325-352); lane 5, Monoclonal antibody (Mab-245), and the immunoprecipitates were analysed by western blotting with anti-Rb antibody 0.47^{10,17,18}.

METHODS. The *NcoI*-*Bam*HI fragment of the full-length Rb coding region derived from pB10¹⁰ was inserted into *NcoI*-*Bam*HI sites of pET8c vector¹². The resulting plasmid pETRbc was transformed into host BL21(DE3)pLysS. Transformed *E. coli* cells were grown in LB media containing 100 μ g ml⁻¹ ampicillin and 25 μ g ml⁻¹ chloramphenicol, and induced by 0.4 mM isopropylthiogalactoside (IPTG) when the culture reached an optical density at 600 nm of 0.6-1. After 2-h induction, 1 l of culture was collected and lysed in lysis buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 2 mM EDTA containing 0.5 mM PMSF) by freezing and thawing. The lysate was clarified by centrifugation at 35,000 r.p.m. for 3 h and the supernatant was fractionated on a 1 $\times$ 15-cm CM-Sepharose CL6B column (Pharmacia) by salt gradient elution (0.1-0.5 M KCl). The fractions enriched to about 80% purity were collected and used for subsequent experiments. Immunological characterization of partially purified p56-Rb was done as previously described¹⁸.

deletion or insertion mutants of p56-Rb were constructed as previously described¹⁰ and expressed in *E. coli* using the T7 expression system (Fig. 3b). Crude bacterial extracts were prepared, and the levels of p56-Rb derivatives in each lysate were determined by western blotting analysis. To examine the T-binding properties of these Rb mutants, bacterial lysates containing equivalent amounts of Rb were mixed with radiolabelled COS cell lysates and the mixture was immunoprecipitated with anti-Rb antibody. Although T antigen was coprecipitated with endogenous Rb in COS cells (Fig. 3a, lane 1), addition of

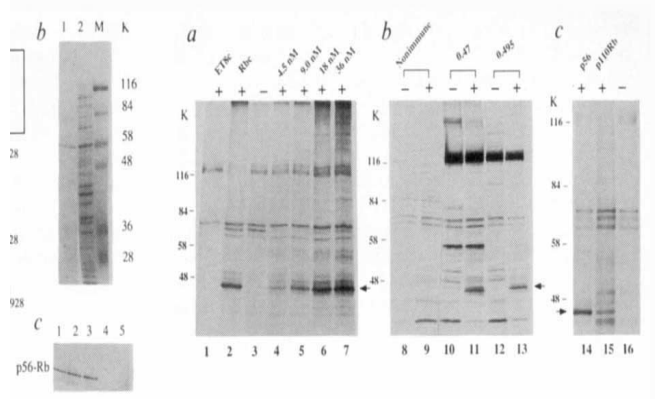
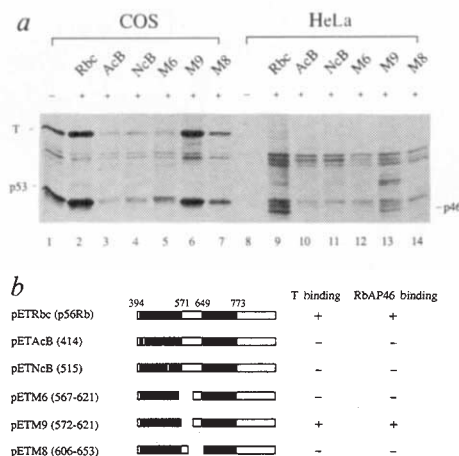


FIG. 2 Detection of RbAP46 in HeLa cell extracts. **a**, Radiolabelled HeLa cell lysates were mixed with the following samples before immunoprecipitation with anti-Rb 0.495: lane 1, control bacterial lysate from IPTG-induced pETRbc transformants; lane 2, lysate from IPTG-induced pETRbc transformants; lane 3, buffer alone; lanes 4-7, increasing amounts of purified p56-Rb as indicated. **b**, HeLa cell lysates were mixed with 0.5 μ g purified p56-Rb (lanes 9, 11 and 13) or buffer alone (lanes 8, 10 and 12), and immunoprecipitated with nonimmune serum (lanes 8 and 9), anti-Rb 0.47 (lanes 10 and 11), and anti-Rb 0.495 (lanes 12 and 13). **c**, HeLa cell lysates were mixed with 0.5 μ g either purified p56-Rb (lane 14) or purified p110-Rb (lane 15) or buffer alone (lane 16) and immunoprecipitated by antibody 0.495. Immunoprecipitates were analysed by 10% SDS-PAGE. Molecular weight calibrations (in K) are shown to the left. Arrows point to the 46K protein.

METHODS. Cells grown nearly confluent were labelled with [³⁵S] methionine for 3 h, washed with cold PBS and then lysed in 1 ml EBC buffer (50 mM Tris-HCl, pH 8.0, 120 mM NaCl, 0.5% Nonidet P-40) containing 200 U ml⁻¹ aprotinin and leupeptin and 1 mM phenylmethylsulphonyl fluoride. The lysates were clarified by centrifugation and aliquots (0.5 ml) of supernatants were taken. Different Rb proteins or control samples were then mixed with each 0.5 ml lysate for 90 min on ice and then incubated with 2 μ g anti-Rb or control antibody for 40 min. Immunocomplexes were collected by adding 80 μ l protein A-Sepharose beads in TBS-BSA buffer for 40 min. After washing five times with 1 ml NET-N buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5% NP-40), the beads were boiled in SDS sample buffer and analysed by SDS-PAGE and fluorography.

wild-type p56-Rb or mutant pETM9 were able to bring down more T antigen than that associated with endogenous Rb (Fig. 3a, compare lanes 2 and 6 with lane 1). Addition of four other mutant proteins actually reduced the amount of coimmunoprecipitating T antigen because the level of anti-Rb antibody was limiting in this analysis. These results essentially replicated our previous definition of the T-binding domains of Rb protein^{9,10}. Interestingly, when the same set of the mutant Rb proteins were added to radiolabelled HeLa cell lysates, the 46K protein coprecipitated with wild-type p56-Rb and pETM9,

FIG. 3 Defining regions of Rb required for association with T and RbAP46. **a**, Radiolabelled COS or HeLa cell lysates were mixed with bacterial lysates containing equivalent amounts of p56-Rb derivatives and immunoprecipitated with anti-Rb 0.47. In COS cells, p53 appeared because it is known to be associated with T, and RbAP46 is not seen because its binding is blocked by excess T. **b**, Structure of the various mutant p56-Rb polypeptides. A schematic representation of p56-Rb is shown on top of the mutant polypeptides. The N-terminal boundary of p56-Rb is not precisely determined yet. **METHODS.** Plasmids pETAcB and pETNcB were constructed by inserting the small *MluI*-*HpaI* fragments derived from digestion of pAcB and pNcB¹⁰ respectively into the corresponding region of pETRbc. Similarly, pETM6, pETM8 and pETM9 were constructed by inserting the small fragments derived from *HpaI* and *DraIII* digestions of plasmids pM6, pM8 and pM9¹⁰ into pETRbc at its corresponding sites. Transformants were grown in 10-ml culture and collected 2 h after induction with isopropylthiogalactoside. Cells were lysed in 1 ml lysis buffer by freezing and thawing, and lysates were spun in a microfuge to clear cell debris. The supernatant was used directly for the mixing experiments.



but not or very weakly with the other four mutants (Fig. 3a, lanes 8–14). These results indicated that the regions of Rb involved in RbAP46 binding coincide precisely with the two regions necessary for binding to T.

If the 46K protein indeed binds to the same domain of Rb as does T, then T antigen should compete effectively with RbAP46 for binding to Rb. In addition, T antigen should be replaceable in competition studies by T peptide (a synthetic 18-residue polypeptide containing amino-acid residues 101–118 of T antigen which are necessary and sufficient for Rb-binding) but not by mutant K peptide which contains the same sequence but with a Lys → Glu substitution at position 107 (ref. 14). This experiment was performed by mixing HeLa cell lysates with p56-Rb, alone or together with T antigen or peptides. Effective competition was observed with 40 nM T antigen (Fig. 4a, lane 4) or 50 μ M T peptide (Fig. 4b, lane 10), although identical amounts of unrelated proteins (data not shown) or mutant K peptide (Fig. 4b, lane 8) were less potent. These results further supported the conclusion that RbAP46 binds to Rb at its T antigen binding domain.

In short, RbAP46 shares the properties and features of Rb binding with T antigen and E1A, oncoproteins of two DNA tumour viruses. Beyond this, the function of the 46K protein and of its association with Rb protein are obscure. It has been proposed that oncoproteins of DNA tumour viruses may transform cells by binding to and inactivating Rb protein. Analogously, one can imagine that the 46K protein may serve as a regulator of Rb to neutralize its biological function under certain cellular conditions. Alternatively, the 46K protein may be a downstream mediator of Rb's normal physiological function. In either case, the lack of RbAP46-Rb complex formation might result in loss of growth control. Consistent with this notion, the regions of Rb involved in T- or RbAP46-binding are common sites for naturally occurring mutations in human tumour cells¹⁰. For example, we have found that a single amino-acid substitution in exon 21, as in small-cell lung tumour cell line H209¹⁵, abolished binding to both T and RbAP46 (data not shown). Although the biological function of RbAP46 has yet to be defined, the inverse correlation of RbAP46 binding to, and biologically significant mutations of, the Rb gene product strongly suggests that the interaction with RbAP46 plays a pivotal role in normal Rb function.

It is not known whether RbAP46 contacts Rb directly, or if intermediate proteins are involved. RbAP46 seems to be a major protein bound to Rb detected by our method, although other

protein species have been inconsistently seen. Perhaps different members of Rb protein can interact with different sets of cellular proteins, depending on cellular growth status. Recently, transcription of the *c-fos* gene has been shown to be downregulated when cells were transfected with Rb¹⁶. It is plausible that Rb-associated protein complexes could bind directly to the *c-fos* promoter, and RbAP46 is a candidate for membership in such complexes. Further characterization of RbAP46 will certainly replace these provocative speculations with real insight into the molecular mechanisms of tumour suppression. □

Received 20 September; accepted 21 December 1990.

1. Klein, G. *Science* **238**, 1539–1545 (1987).
2. Huang, H.-J. S. et al. *Science* **242**, 1563–1566 (1988).
3. Bookstein, R., Shew, J.-Y., Chen, P.-L., Scully, P. & Lee, W.-H. *Science* **247**, 712–715 (1990).
4. Sumegi, J., Uzvolgyi, E. & Klein, G. *Cell Growth Differ.* **1**, 247–250 (1990).
5. Lee, W.-H., Bookstein, R. & Lee, E. Y.-H. P. in *Tumor Suppressor Genes* (ed. Klein, G.) 169–200 (Marcel Dekker, New York, 1990).
6. Whyte, P. et al. *Nature* **334**, 124–129 (1988).
7. DeCaprio, J. A. et al. *Cell* **54**, 275–283 (1988).
8. Dyson, N., Howley, P. M., Munger, K. & Harlow, E. *Science* **243**, 934–937 (1989).
9. Hu, Q., Dyson, M. N. & Harlow, E. *EMBO J.* **9**, 1147–1155 (1990).
10. Huang, S., Wang, N.-P., Tseng, B. Y., Lee, W.-H. & Lee, Y.-H. P. *EMBO J.* **9**, 1815–1822 (1990).
11. Xu, H.-J., Hu, S.-X., Hashimoto, T., Takashi, R. & Benedict, W. F. *Oncogene* **4**, 807–812 (1989).
12. Rosenberg, A. H. et al. *Gene* **56**, 125–135 (1987).
13. Wang, N.-P., Qian, Y.-W., Chung, A. E., Lee, W.-H. & Lee, E. Y.-H. P. *Cell Growth Differ.* **1**, 429–437 (1990).
14. DeCaprio, J. A. et al. *Cell* **58**, 1085–1095 (1989).
15. Bignon, Y.-J. et al. *Cell Growth Differ.* **1**, 647–651 (1990).
16. Robbins, P. D., Horowitz, J. M. & Mulligan, R. C. *Nature* **346**, 668–671 (1990).
17. Wang, N.-P. et al. *Cell Growth Differ.* **1**, 233–239 (1990).
18. Shew, J.-Y., Ling, N., Yang, X., Fodstad, O. & Lee, W.-H. *Oncogene Res.* **1**, 205–214 (1989).

ACKNOWLEDGEMENTS. We thank F. W. Studier for plasmids and strains of the T7 expression system, and F. Hong, C. Lin, M. Kapiloff and R. Chen for discussions concerning the use of the system; N.-P. Wang and Y.-W. Qian for p110-Rb from insect cells, and for anti-Rb antibodies 0.495 and 0.47; B. Tseng for SV40 T antigen; N. Lin for peptide synthesis and purification; and R. Bookstein for reading the manuscript. This work was supported by the NIH and Council for Tobacco Research to E. Y.-H.P.L. and W.-H.L. S.H. was supported in part by CRCC, University of California, and by a Fight For Sight—Prevent Blindness Postdoctoral Fellowship.

Alternative temporal control systems for hypodermal cell differentiation in *Caenorhabditis elegans*

Zhongchi Liu & Victor Ambros

Department of Cellular and Developmental Biology, Harvard University, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA

IN certain multicellular organisms, genetic regulatory systems that specify the timing of cell division, differentiation and morphogenesis^{1–3} must accommodate environmental and physiological contingencies that perturb or arrest development. For example, *Caenorhabditis elegans* can either develop continuously through four larval stages (L1–L4) or arrest indefinitely as a 'dauer larva' at the second larval (L2) moult, and later resume L3 and L4 development^{4–7}. At the larva-to-adult (L4) moult of both continuous and 'post-dauer' development, hypodermal cells switch (the 'L/A switch') from a proliferating state to the terminally differentiated state. Four temporal regulators, *lin-4*, *lin-14*, *lin-28* and *lin-29*, have been identified in *C. elegans* by mutations that cause precocious or retarded expression of stage-specific post-embryonic development events, including the L/A switch (refs 3, 8, 9; Fig. 1a). These genes have been organized into a genetic pathway that controls the timing of the L/A switch during continuous development¹⁰: *lin-29* activates the switch and the other heterochronic genes regulate it indirectly by regulating *lin-29*. We have now examined how the proper timing of this event is specified in alternative developmental pathways. In continuously developing *lin-4*, *lin-14* and *lin-28* mutants the L/A switch occurs at abnormally early or late moults^{3,8}, but during post-dauer development of the same mutants the L/A switch occurs normally. Thus

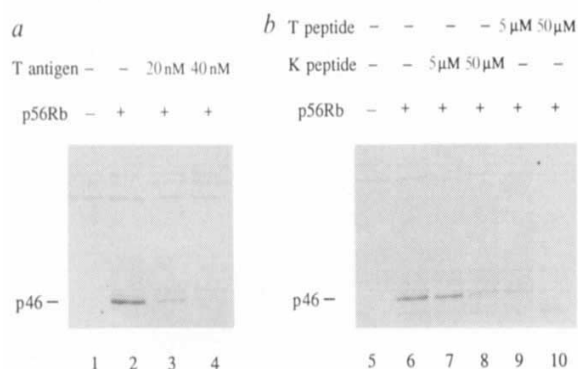


FIG. 4 T antigen and T peptide compete with RbAP46 for binding to Rb. Radiolabelled HeLa cell lysates were mixed with different samples as indicated on top of each lane and incubated for 90 min on ice. Immunoprecipitation was done by anti-Rb 0.495 and subsequently analysed as described in Fig. 2. The amino-acid sequence (one-letter code) of T peptide is ENLFCSEEMPSSDDEAAT, and the K peptide is ENLFCSEEMPSSDDEAAT. These two peptides differ only in one amino acid at position 7. About 20 nM of p56-Rb was used where indicated.

TABLE 1 Expression of the L/A switch in heterochronic mutants during continuous and post-dauer development

Genotype	Heterochronic expression of the L/A switch			
	Continuous	Precocious	Post-dauer	Retarded
Wild type	—	—	—	—
<i>lin-14</i> <i>n355 n679ts</i> 15 °C	—	—	100% (21/21)	19%* (4/21)
<i>n536/+</i>	—	—	100% (<i>n</i> > 100)	3%* (2/73)
<i>n355 n679ts</i> 25 °C	100% (40/40)	0% (0/36)	—	—
<i>n727</i>	100% (55/55)	0% (0/51)	—	—
<i>n360</i>	100% (<i>n</i> > 100)	0% (0/28)	—	—
<i>ma135</i>	100% (<i>n</i> > 100)	0% (0/39)	—	—
<i>n179ts</i> 20 °C	0% (<i>n</i> > 100)	0% (<i>n</i> > 100)	—	—
<i>lin-4(e912), lin-14(n179ts)</i> 20 °C	—	—	100% (9/9)	23% (8/35)
<i>lin-28</i> <i>n1119</i>	100% (33/33)	3%† (1/31)	—	—
<i>n719</i>	100% (<i>n</i> > 100)	8%† (3/37)	—	—
<i>lin-29</i> <i>n333</i>	—	—	100% (<i>n</i> > 100)	100% (55/55)
<i>n1440</i>	—	—	100% (<i>n</i> > 100)	100% (<i>n</i> > 100)

During continuous development, loss-of-function (*lf*) mutations of *lin-14* and *lin-28* cause the L/A switch to occur precociously (refs 3, 10; Fig. 1a). 'Precocious' L/A switch refers to the formation of adult lateral alae and absence of seam cell division at the L3 moult. In contrast, *lin-4(e912)*, *lin-29(lf)* and the semi-dominant gain-of-function (*gf*) mutations of *lin-14(n536)* delay or completely inhibit the L/A switch (refs 3, 8, 9; Fig. 1a); their seam cells synthesize a larval cuticle and continue cell divisions at the L4 moult and they develop through supernumerary larval moults. 'Retarded' L/A switch refers to absence of lateral alae and occurrence of seam cell division at the L4 moult. The percentage of animals that express a precocious or retarded L/A switch is given. Where known precisely, number of animals examined is indicated in parentheses; '*n* > 100' is an estimation of the number of animals examined. — Corresponding mutant phenotype was never observed. Seam cell divisions were observed using Nomarski optics¹⁸. Adult lateral alae were visualized by side illumination under $\times 50$ magnification in the dissecting microscope³. The *lin-14(n355n679ts)* (strain MT1388) was derived by reverting the semi-dominant gain-of-function *lin-14(n355)* mutation, and thus contains both the original semi-dominant mutation *n355* and the new temperature-sensitive intragenic suppressor *n679ts*. *lin-14(n355n679ts)* at 15 °C is semi-dominant and results in retarded defects. At 25 °C, *lin-14(n355n679ts)* exhibits precocious defects⁹. As severely retarded mutants such as *lin-14(n536)* and *lin-4(e912)* do not form dauer larvae, we used less retarded strains, *lin-14(n536/+)* and *lin-4(e912); lin-14(n179ts)*. In the *lin-4(e912); lin-14(n179ts)* double mutant at 20 °C, the retarded L/A switch is likely to be solely caused by the *lin-4(e912)* mutation, because *lin-14(n179ts)* animals at 20 °C are weakly precocious.

* Animals expressed the precocious or the retarded L/A switch in only local area(s) of the body after post-dauer development (see text).

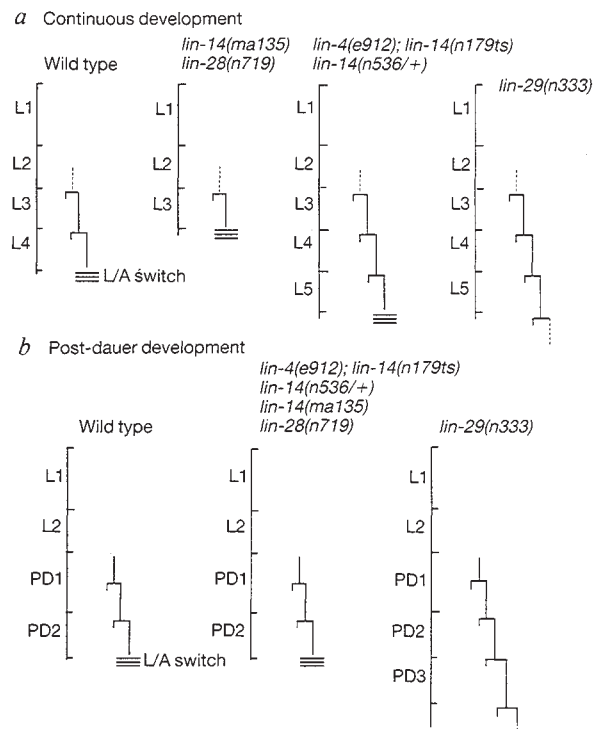
† Occasionally, *lin-28* mutants displayed very small patches (corresponding to 1 or 2 seam cells) of precocious adult cuticle even after post-dauer development. This may result from the fact that some *lin-28* animals express patches of adult cuticle at the L2-to-dauer moult (data not shown), and thus these cells may have terminally differentiated before dauer larva formation.

hypodermal cell differentiation is regulated by separate temporal control systems, depending on the developmental history.

Dauer larvae were isolated from starved cultures by SDS selection⁴, placed on food, and post-dauer development was examined. The morphological and behavioural defects that result from abnormal timing of the L/A switch in both retarded and precocious mutants of *lin-14*, *lin-4* and *lin-28* were efficiently

suppressed in animals that developed from dauer larvae (Table 1). Cell-lineage analysis revealed that the precocious or retarded L/A switch defects were suppressed by the expression of a wild-type post-dauer developmental programme consisting of two larval stages (PD1 and PD2, analogous to L3 and L4) with the L/A switch at the PD2-to-adult moult (Fig. 1b). Suppression absolutely required dauer larva formation: mutant animals that

FIG. 1 Lateral hypodermal cell lineages during continuous and post-dauer development of wild type and heterochronic mutants. The vertical axis indicates the four post-embryonic larval stages, with horizontal marks indicating the times of hatching and each larval moult. L/A switch is indicated by triple bars. Cell lineages were followed in living nematodes at 20 °C using Nomarski optics¹⁸. a, Representative lateral hypodermal cell lineages, beginning at the L2 moult of continuous development (previous cell divisions of this lineage are not shown) of the wild type, two precocious mutants (*lin-14(ma135)* and *lin-28(n719)*) and three retarded mutants (*lin-4(e912); lin-14(n179ts)*, *lin-14(n536/+)* and *lin-29(n333)*). b, Representative lateral hypodermal cell lineages during post-dauer development in precocious and retarded mutants beginning at the time of feeding developmentally arrested dauer larvae. PD1 and PD2 refer to the two post-dauer larval stages. Dauer larva arrest occurred at the second (L2) moult in all these animals. The diagrams are composites of lineages observed in several animals of each genotype. Some post-dauer lineages were followed intermittently¹⁰; the results of intermittent observation of multiple animals were confirmed by continuous observation of one or more individuals.



did not become dauer larvae in starved and crowded cultures were not suppressed, but dauer larvae from the same cultures were all suppressed. Suppression was not a function of time spent in dauer larva arrest, and dauer larvae induced by growth in the presence of the dauer-promoting pheromone⁶, by starvation, or by the *ts* dauer-constitutive mutations¹¹ were equally suppressed. Finally, the mechanism of suppression does not result from heritable suppressor mutations or a restoration of *lin-4*, *lin-28* or *lin-14* gene activity. Presumptive null alleles, such as *lin-14(ma135)* and *lin-28(n719)* (ref. 10), show essentially full suppression. Thus the post-dauer developmental programme does not require *lin-4*, *lin-14* or *lin-28*.

The *lin-4* and *lin-14* mutants show an essentially wild-type post-dauer developmental programme even when they begin post-dauer development at abnormal stages. Although wild-type animals form dauer larvae only at the L2 moult, certain heterochronic mutants sometimes arrest as dauer larvae at abnormal stages—at the L1 moult in precocious *lin-14* mutants, or at the L3 moult in retarded *lin-14* and *lin-4* mutants¹². The L1 or L3 dauer larvae execute an essentially normal post-dauer pathway, consisting of two larval stages and two rounds of lateral hypo-

dermal cell division, with the L/A switch at the end of the PD2 stage (data not shown). Thus the post-dauer programme is essentially independent of events before dauer larva arrest. Interestingly, the PD1 stage was lengthened by 10–15 h (about the length of one larval stage) for L1 dauer larvae and was shortened by about 5–6 h for L3 dauer larvae. This suggests that the PD1 may be naturally flexible in length to allow for the resynchronization of development between the hypodermis and other tissues.

Generally, post-dauer development leads to complete suppression of heterochronic L/A switch defects (Table 1). However, certain mutants with elevated *lin-14* activity (*lin-14(gf)* or *lin-4(e912)* (ref. 10)) that display retarded defects after continuous development also display a residual retarded L/A-switch defect after post-dauer development (Table 1). In animals whose post-dauer development was followed using Nomarski optics, we found that this residual retarded defect in patches of cells resulted from 'incomplete dauer larva morphogenesis'¹² in the corresponding cell lineages. This indicates that expression of the post-dauer programme is variably blocked by elevated *lin-14* activity, and can occur cell-intrinsically. In contrast to these retarded mutants, essentially all precocious mutant animals were completely suppressed, although precocious mutants readily form incomplete dauer larvae (ref. 12, and Table 1). This indicates that expression of the post-dauer developmental programme in a particular hypodermal lineage is not sensitive to reduced *lin-14* or *lin-28* activity and does not require prior dauer larva differentiation.

The *lin-29* gene is unique among the heterochronic genes examined so far in that its loss causes a complete block in the execution of the L/A switch during post-dauer development as well as in continuous development (Table 1; Fig. 1). The *lin-29* gene is thought to be a direct activator of the L/A switch, and *lin-4*, *lin-14* and *lin-28* to regulate the stage-specific activation of *lin-29* (ref. 10). Our observations that post-dauer development suppresses the L/A switch defects of *lin-4*, *lin-14* and *lin-28* mutants but not of *lin-29* mutants suggests that *lin-29* is critical for both continuous and post-dauer development, but is subject to modes of regulation specific to each pathway. The only genes known to affect recovery from the dauer larva stage are *daf* formation genes (*daf*) that also control entry into the dauer larva stage¹¹, but none of the *daf* mutants that we have examined display defects in the timing of the L/A switch during post-dauer or continuous development (data not shown).

Certain stage-specific collagen genes are regulated by *lin-29* during continuous and post-dauer development (Z.L. and V.A., unpublished). We tested whether post-dauer development restores the proper temporal specificity of these genes in heterochronic mutants. A *col-19-β-gal* chimaeric gene that is expressed adult-specifically in wild type during both developmental pathways (Fig. 2a, b, e, f; refs 13, 14) is expressed precociously in *lin-14(n179ts)* animals that develop continuously at the restrictive temperature (Fig. 2c, d). This precocious expression is corrected by post-dauer development (Fig. 2g, h). Thus post-dauer suppression of defects in the stage-specificity of cuticle morphogenesis reflects a restoration of the proper stage-specific transcription of *col-19* and perhaps of other genes regulated by *lin-29*. Although *lin-29* encodes a putative transcription factor (A. Rougvié and V.A. personal communication), we do not know if the *lin-29* product regulates *col-19* directly or indirectly.

Certain *C. elegans* mutants defective in cuticle structure or vulva-cell lineage exhibit a weaker or modified phenotype after post-dauer development^{15,16}. These cases of post-dauer partial suppression or modification of phenotype differ quantitatively and qualitatively from the suppression of heterochronic mutants. The suppression is not as efficient as for the heterochronic mutants, and among vulva-cell specification and cuticle structure genes, post-dauer suppression seems to be the exception. In contrast, the entire regulatory pathway consisting of *lin-4*, *lin-14* and *lin-28* is completely dispensable for post-dauer develop-

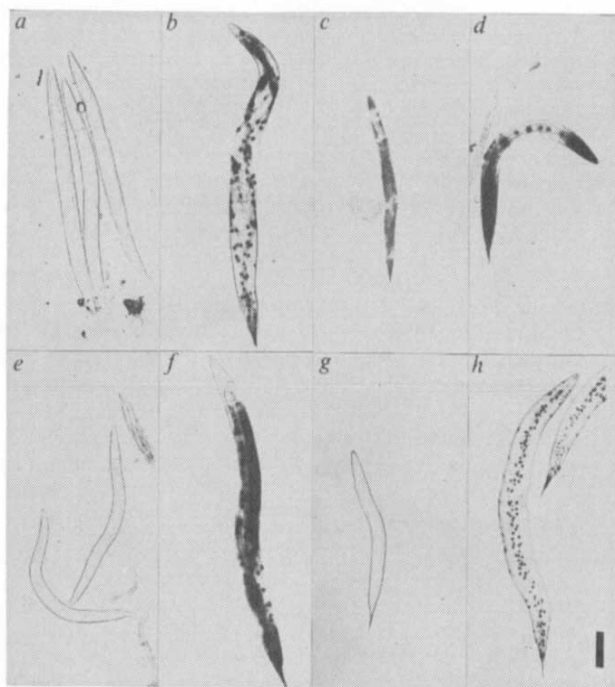


FIG. 2 Expression of a *col-19-β-gal* chimaeric gene in the wild type and *lin-14(n179ts)* mutants during continuous and post-dauer development. Top panel, animals that developed through continuous development: a, wild-type L4 larvae; b, wild-type adult; c, *lin-14(n179ts)* young L4 larva at 25 °C; d, *lin-14(n179ts)* adult at 25 °C. Bottom panel, animals that developed through post-dauer development: e, wild-type L4 larvae; f, wild-type adult; g, *lin-14(n179ts)* young L4 larva at 25 °C; h, *lin-14(n179ts)* adult at 25 °C. Animals were from strains of the indicated genotype stably transformed with the *col-19-β-gal* gene and a dominant marker *rol-6(sul1006)* (ref. 19; C. Mello and V.A., unpublished) by microinjection of DNA into the hermaphrodite gonad. The *col-19-β-gal* chimaeric gene was a translational fusion between *col-19* and *β-galactosidase* gene in vector 22.04 (ref. 20; Z.L. and V.A., unpublished). As the *col-19-β-gal* fusion contains only 7 amino acids of the *col-19* gene, the adult-specific *col-19-β-gal* expression probably reflects adult-specific transcriptional initiation. *β-Galactosidase* activity in transgenic worms was assayed by staining fixed worms with a solution containing X-gal (5-bromo-4-chloro-3-indolyl- α -galactoside) (A. Fire and J. Shaw, personal communication). The *col-19-β-gal* fusion has a nuclear localization signal²⁰ that causes most of the staining to be localized to hypodermal nuclei, although in more intensely stained animals some of the staining apparently leaked from the nuclei into other areas of the body. Scale bar, 100 μ m.

ment. These differences may reflect particularly acute constraints imposed on temporal control systems (as opposed to spatial patterning or cuticle morphogenesis) by the suspension of development at dauer larva arrest.

What mechanistic or developmental constraints account for the existence of separate temporal signals for continuous and post-dauer development? Heterochronic genes may encode components of a temporally dynamic system of interacting gene products. The *lin-14* activity must decrease between L1 and L2 to specify the proper timing and sequence of diverse events, including the L/A switch^{9,10,17}; *lin-4* and *lin-28*, which interact with *lin-14*, might also experience precise temporal changes. The indefinite period of developmental arrest in the dauer larva stage would inevitably interrupt the dynamics of this system, and thus render *lin-4*, *lin-14* and *lin-28* useless for conveying temporal information to cells during post-dauer development. Other organisms with optional developmentally arrested stages, such as insects that undergo diapause, may use a similar strategy to cope with the temporal control problems necessarily associated with alternative developmental histories. □

Received 5 December; accepted 27 December 1990.

1. Alberch, P., Gould, S. J., Oster, G. F. & Wake, D. B. *Paleobiology* **5**, 296–317 (1979).
2. Gould, S. J. *Ontogeny and Phylogeny* (Harvard University Press, Cambridge, Massachusetts, 1977).
3. Ambros, V. & Horvitz, H. R. *Science* **226**, 409–416 (1984).
4. Cassada, R. C. & Russell, R. L. *Dev Biol.* **46**, 326–342 (1975).
5. Evans, A. A. F. & Perry, R. M. in *The Organization of Nematodes* (ed. Croll, N. A.) (Academic, New York, 1976).
6. Golden, J. M. & Riddle, D. L. *Science* **218**, 578–580 (1987).
7. Riddle, D. L. in *The Nematode Caenorhabditis elegans* (eds Wood, W. B. and the community of *C. elegans* researchers) (Cold Spring Harbor Laboratory, New York, 1988).
8. Chalfie, M., Horvitz, R. H. & Sulston, J. E. *Cell* **24**, 59–69 (1981).
9. Ambros, V. & Horvitz, H. R. *Genes Dev.* **1**, 398–414 (1987).
10. Ambros, V. *Cell* **57**, 49–57 (1989).
11. Riddle, D. L., Swanson, M. M. & Alberts, P. S. *Nature* **290**, 668–671 (1981).
12. Liu, Z. & Ambros, V. *Genes Dev.* **3**, 2039–2049 (1989).
13. Cox, G. N., Fields, C., Kramer, J. M., Rosenzweig, B. & Hirsh, D. *Gene* **76**, 331–344 (1989).
14. Cox, G. N. & Hirsh, D. *Molec. cell. Biol.* **5**, 363–372 (1985).
15. Ferguson, E. & Horvitz, H. R. *Genetics* **110**, 17–72 (1985).
16. Cox, G. N., Laufer, J. S., Kusch, M. & Edgar, R. *Genetics* **95**, 317–339 (1980).
17. Ruvkun, G. & Guisto, J. *Nature* **338**, 313–319 (1989).
18. Sulston, J. E. & Horvitz, R. H. *Dev Biol.* **56**, 110–156 (1977).
19. Kramer, J. M., French, R. P., Park, E. & Johnson, J. J. *Molec. cell. Biol.* **10** (1990).
20. Fire, A., Harrison, S. & Dixon, D. *Gene* **93**, 189–198 (1990).

ACKNOWLEDGEMENTS. We thank S. Euling, R. Feinbaum, G. Struhl, D. Levitan, C. Mello, A. Rougvié and other members of the laboratory for critical reading of the manuscript, and P. Cherbas for comments.

Function of DnaJ and DnaK as chaperones in origin-specific DNA binding by RepA

Sue Wickner*, Joel Hoskins† & Keith McKenney†

* Laboratory of Molecular Biology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA

† Center for Advanced Research in Biotechnology, National Institute of Standards and Technology, Rockville, Maryland 20850, USA

HEAT-shock proteins are normal constituents of cells whose synthesis is increased on exposure to various forms of stress. They are interesting because of their ubiquity and high conservation during evolution. Two families of heat-shock proteins, hsp60s and hsp70s, have been implicated in accelerating protein folding and oligomerization and also in maintaining proteins in an unfolded state, thus facilitating membrane transport^{1–5}. The *Escherichia coli* hsp70 analogue, DnaK, and two other heat-shock proteins, DnaJ and GrpE, are required for cell viability at high temperatures and are involved in DNA replication of phage λ and plasmids P1 and F^{6–10}. These three proteins are involved in replication *in vitro* of P1 DNA along with many host replication proteins and the P1 RepA initiator protein^{11,12}. RepA exists in a stable protein complex with DnaJ containing a dimer each of RepA and DnaJ¹¹. We

TABLE 1 Temperature requirement for activation of RepA

Incubation at 24 °C (15 min)	Incubation at 0 °C (15 min)	<i>oriP1</i> DNA bound (%)
Complete	No additions	26
No additions	Complete	<2
Minus ATP	+ATP	<2
Minus <i>oriP1</i> DNA	+ <i>oriP1</i> DNA	32
Minus DnaJ	+DnaJ	4
Minus DnaK	+DnaK	4
Minus RepA	+RepA	3

DNA binding was measured by retention on nitrocellulose filters. Reaction mixtures for the first incubation at 24 °C contained the components described in Fig. 1 with the omissions indicated and with 20 ng RepA, 30 ng DnaJ and 100 ng DnaK. After 15 min, the reactions were chilled on ice, the omitted components were added, and the reactions were continued for 15 min at 0 °C.

report here that DnaK and DnaJ mediate an alteration in the P1 initiator protein, rendering it much more active for *oriP1* DNA binding.

RepA binds specifically to five 19-base-pair (bp) direct repeats in the P1 origin¹³. We observed that the DnaJ–RepA protein complex also binds to the P1 origin (Fig. 1a). Surprisingly, about a 100-fold molar excess of RepA dimers to binding sites was required for binding by RepA alone or in a complex with DnaJ (Fig. 1a). These results suggested that the preparations of RepA and DnaJ–RepA complex were largely inactive or the affinity of RepA for DNA was very low.

Because it seemed likely that the other heat-shock proteins may function with DnaJ, we added DnaJ, DnaK and GrpE to DNA-binding reactions with RepA. About 100-fold less RepA was required for *oriP1* DNA binding when DnaJ, DnaK and ATP were added to the reaction mixtures (Fig. 1a–c). GrpE was not required for RepA activation (Fig. 1c). DnaJ–RepA protein complexes could be activated by DnaK and ATP, without additional DnaJ. ATP hydrolysis is probably required because the non-hydrolysable ATP analogue ATP- γ -S was inactive and was a competitive inhibitor of the reaction.

The stimulation by DnaJ, DnaK and ATP of *oriP1* DNA binding by RepA depended both on the time of incubation and the temperature. DNA binding increased linearly at 24 °C for 20 min, but was barely detectable when the reactions were carried out at 0 °C (Fig. 2). Because protein–DNA binding is expected to occur at 0 °C, we attempted to divide the DnaJ- and DnaK-stimulated binding reaction. In a first reaction at 24 °C, we incubated all five components (DnaJ, DnaK, RepA, ATP and *oriP1* DNA) or four of them. Then in a second incubation at 0 °C, we added the omitted component and measured *oriP1* DNA bound (Table 1). DnaK, DnaJ, RepA and ATP were each required for DNA binding in the first incubation, but *oriP1* DNA was not. The first reaction in the absence of DNA proceeded linearly for 20 min at 24 °C, whereas the DNA-binding reaction was complete by 30 s at 0 °C (Fig. 2). The same requirements were necessary when the first reaction was carried out at 30 °C or 42 °C. When DnaJ and RepA were replaced by the DnaJ–RepA complex, the activation reaction still required a 20-min incubation at 24 °C with DnaK and ATP. This is consistent with our previous observation that the DnaJ–RepA complex forms rapidly at 0 °C on mixing the two proteins¹¹. These experiments show that DnaJ and DnaK, in an ATP and temperature-dependent reaction, activate RepA in the absence of DNA so that it can then rapidly bind to *oriP1* DNA.

Of the three proteins required for the reaction, only RepA ends up bound to the P1 DNA. We prepared a large reaction mixture, isolated the protein–DNA complex by gel filtration, subjected the material to SDS-PAGE, and stained for protein. Densitometer analysis of the gel showed that RepA was the only protein detectable (Fig. 3). Protein–DNA complexes similarly

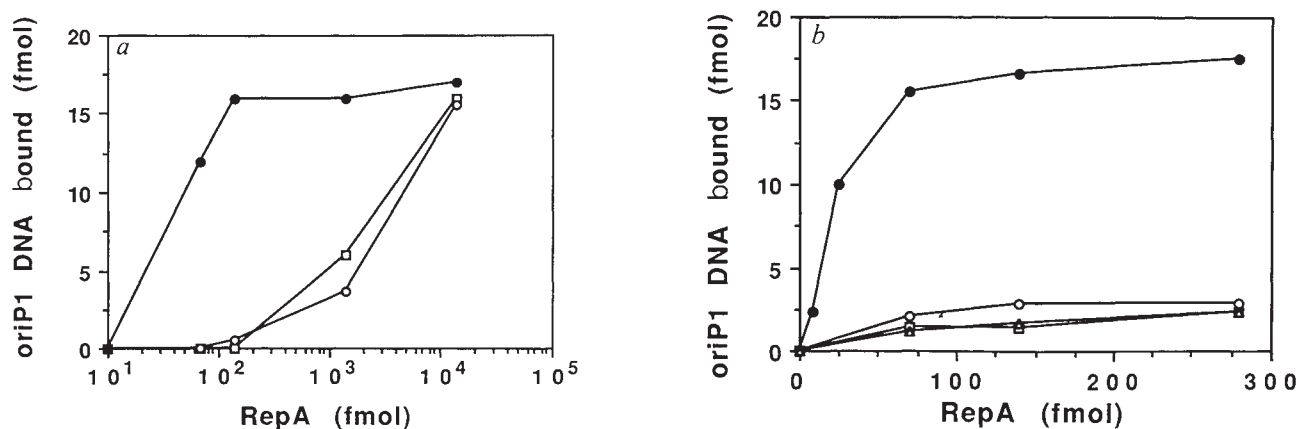
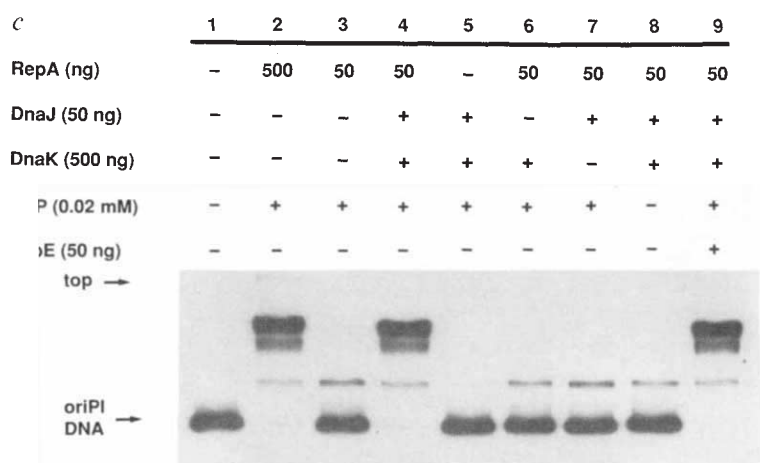


FIG. 1 *a*, oriP1 DNA binding by RepA alone (○), DnaJ-RepA complex alone (□) and RepA in combination with 30 ng DnaJ, 100 ng DnaK and ATP (●). The molar amount of RepA is calculated from the observations that RepA is a dimer in solution^{11,13} and the DnaJ-RepA complex contains a dimer each of RepA and DnaJ¹¹. *b*, Requirements for activation of RepA for oriP1 DNA binding. ●, Complete reactions containing RepA, 30 ng DnaJ and 100 ng DnaK; reactions lacking ATP (○), DnaK (□) or DnaJ (△). *c*, oriP1 DNA binding as measured by gel retardation. DNA binding reactions were carried out as described for above but with 25 μ M ATP, 5 μ g ml⁻¹ calf thymus DNA and proteins in amounts indicated. After 20 min at 24 °C, reactions were analysed by electrophoresis on 5% acrylamide gels¹³. Intermediate shifted bands were generated by subsaturating amounts of RepA¹³.

METHODS. Reaction mixtures (20 μ l) contained 20 mM Tris-HCl, pH 7.5, 40 mM KCl, 100 mM NaCl, 10 mM MgCl₂, 0.1 mM EDTA, 1 mM dithiothreitol, 50 μ M ATP, 50 μ g ml⁻¹ BSA, 50 μ g ml⁻¹ calf thymus DNA, 5 ng labelled oriP1 DNA (285-bp fragment containing P1 origin DNA from position 381 to 611¹⁶), DnaK, DnaJ and RepA. Purified DnaK, DnaJ, GrpE and RepA proteins¹¹ were more than 95% pure as determined by SDS-PAGE; DnaJ-RepA complex isolated from cell extracts¹¹ was 80% pure. Incubations were for 20 min at 24 °C. Reaction mixtures were then



filtered through nitrocellulose filters and filter-bound radioactivity was measured.

isolated from reaction mixtures lacking either DnaJ or ATP contained no detectable proteins. Thus, DnaJ and DnaK activate RepA to bind to oriP1 DNA but are not part of the protein-DNA complex.

The molecular change in RepA brought about by DnaJ and DnaK is not so far understood. Activation does not seem to involve proteolytic processing, because activated RepA isolated from DNA or from reaction mixtures comigrated with purified RepA on SDS-PAGE. DnaK autophosphorylates¹⁴, but there was no detectable phosphorylation of RepA during the activation reaction. DnaK and other hsp70s facilitate disaggregation^{4,5}, but RepA does not appear to aggregate as measured by gel filtration of reaction mixtures containing RepA alone or

RepA plus DnaJ and DnaK but not ATP. DnaJ and DnaK could dissociate RepA dimers to monomers or cause some subtle conformational change in RepA.

DnaJ, DnaK and GrpE may also function *in vivo* to activate DNA binding by RepA. Repression of the autoregulated *repA* promoter by RepA is 2- to 10-fold less in *dnaJ*, *dnaK* and *grpE* mutant strains compared to wild-type strains¹⁵. This could be explained by our finding that DnaJ and DnaK activate the DNA-binding function of RepA. And what is the biochemical function of GrpE? We have demonstrated that it is required for P1 replication *in vitro* but that it is not required for the activation of RepA.

A model for RepA activation is shown in Fig. 4. The inactive

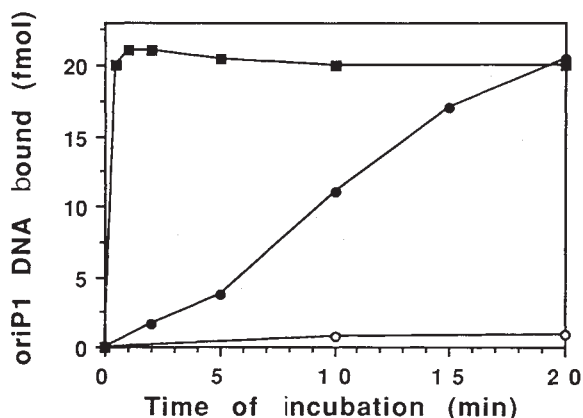


FIG. 2 Time course of RepA activation. DNA binding was measured by retention on nitrocellulose filters as described in Fig. 1 with 30 ng RepA, 30 ng DnaJ and 200 ng DnaK. Reactions were incubated for the times indicated at 0 °C (○) or at 24 °C (●). Other reactions were preincubated for 20 min at 24 °C in the absence of DNA. Then oriP1 DNA was added and the reactions were continued for the times indicated at 0 °C (■).

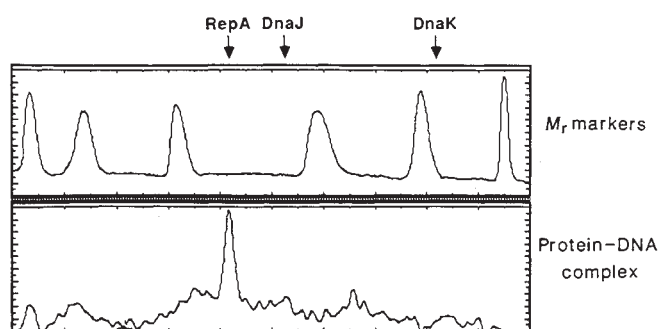


FIG. 3 Determination of proteins associated with *oriP1* DNA. The reaction was as described in Fig. 1, but in 2 ml containing 5 μ g RepA, 5 μ g DnaJ, 50 μ g DnaK, 500 ng [3 H]*oriP1* DNA (1,182-bp DNA fragment) and 5 μ g ml $^{-1}$ calf thymus DNA; BSA was omitted. After 20 min at 24 $^{\circ}$ C, the reaction was applied to a 1 $\times$ 25 cm Sepharose 6B column equilibrated with 20 mM Tris-HCl, pH 7.5, 10 mM MgCl $_2$, 1 mM dithiothreitol, 0.1 mM EDTA, 40 mM KCl and 100 mM NaCl. The column was eluted with the same buffer. Fractions in the void volume (containing radioactive *oriP1* DNA and associated proteins) were concentrated by precipitation with 15% tricarboxylic acid. The pellet was washed with acetone, dissolved in SDS sample buffer, analysed by 10–15% SDS-PAGE, stained with Coomassie blue, and scanned with a densitometer. Molecular weight markers were BioRad low-range SDS-PAGE standards.

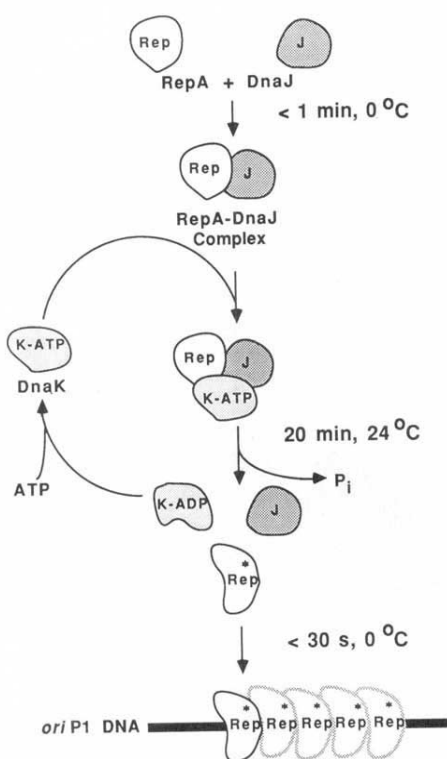


FIG. 4 Model for activation of RepA.

form of RepA binds tightly and rapidly to DnaJ at 0 $^{\circ}$ C to form a 2:2 tetramer 11 . DnaJ may be as a protein cofactor that targets RepA for subsequent recognition and activation by DnaK in an ATP and temperature-dependent reaction. The energy of ATP hydrolysis is probably used to activate RepA and to release DnaJ. The activated RepA then can bind rapidly to DNA at 0 $^{\circ}$ C. This activation may be involved in the regulation of the rate of initiation of DNA replication and the regulation of expression of the autoregulated *repA* gene.

To our knowledge, this is the first demonstration in an *in vitro* system that DnaJ and DnaK function as molecular chaperones

to activate a sequence-specific DNA binding protein. Because the heat-shock proteins are so highly conserved between bacteria and eukaryotes, it is likely that similar mechanisms operate to regulate gene expression and replication in other systems. $\square$

Received 16 November 1990; accepted 10 January 1991.

1. Rothman, J. E. *Cell* **59**, 591–601 (1989).
2. Neidhardt, F. & VanBogelen, R. in *Escherichia coli and Salmonella typhimurium* (eds Neidhardt, F. et al.) 1334–1345 (American Society for Microbiology, Washington, DC, 1987).
3. Lindquist, S. & Craig, E. A. *Rev. Genet.* **22**, 631–677 (1988).
4. Pelham, H. R. B. in *Stress Proteins in Biology and Medicine* (eds Morimoto, R., Tissieres, A. & Georgopoulos, C.) 287–299 (Cold Spring Harbor Laboratory Press, New York, 1990).
5. Morimoto, R. I., Tissieres, A. & Georgopoulos, C. in *Stress Proteins in Biology and Medicine* (eds Morimoto, R., Tissieres, A. & Georgopoulos, C.) 1–36 (Cold Spring Harbor Laboratory Press, New York, 1990).
6. Georgopoulos, C., Ang, D., Liberek, K. & Zyliz, M. in *Stress Proteins in Biology and Medicine* (eds Morimoto, R., Tissieres, A. & Georgopoulos, C.) 191–221 (Cold Spring Harbor Laboratory Press, New York, 1990).
7. Bukau, B. & Walker, G. *J. Bact.* **171**, 6030–6038 (1989).
8. Tilly, K. & Yarmolinsky, M. *J. Bact.* **171**, 6025–6029 (1989).
9. Kawasaki, Y., Wada, C. & Yura, T. *Molec. gen. Genet.* **220**, 277–282 (1990).
10. Ezaki, B., Ogura, T., Mori, H., Niki, H. & Hiraga, S. *Molec. gen. Genet.* **218**, 183–189 (1989).
11. Wickner, S. *Proc. natn. Acad. Sci. U.S.A.* **87**, 2690–2695 (1990).
12. Wickner, S. & Chatteraj, D. *Proc. natn. Acad. Sci. U.S.A.* **84**, 3668–3672 (1987).
13. Abeles, A. L. *J. biol. Chem.* **261**, 3548–3555 (1986).
14. Zyliz, M., LeBowitz, J., McMacken, R. & Georgopoulos, C. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6431–6435 (1983).
15. Tilly, K., Sozhamannan, S. & Yarmolinsky, M. *The New Biologist* **2**, 1–6 (1990).
16. Abeles, A. L., Snyder, K. M. & Chatteraj, D. K. *J. molec. Biol.* **173**, 307–324 (1984).

ACKNOWLEDGEMENTS. We thank Nic Cozzarelli and Mike Maurizi for their helpful discussions and comments on this manuscript.

Structural architecture of an outer membrane channel as determined by electron crystallography

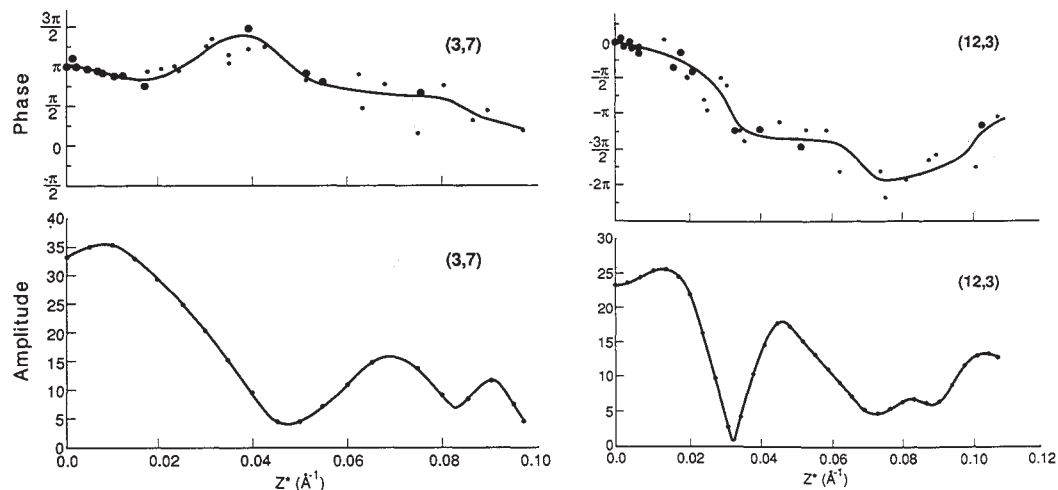
B. K. Jap, P. J. Walian & K. Gehring

Donner Laboratory, Lawrence Berkeley Laboratory, University of California, Berkeley, California 94720, USA

PORINS are a family of membrane channels commonly found in the outer membranes of Gram-negative bacteria where they serve as diffusional pathways for waste products, nutrients and antibiotics, and can also be receptors for bacteriophages 1,2 . Porin channels have been shown *in vitro* to be voltage-gated $^{3-6}$. They can exhibit slight selectivities for certain solutes; for example PhoE porin has some selectivity for anionic and phosphate-containing compounds 1,7 . Unlike many known membrane proteins which often contain long stretches of hydrophobic segments that are believed to traverse the membrane in a helical conformation, porins are found to have charged residues distributed almost uniformly along their primary sequences and have most of their secondary structure in a β -sheet conformation $^{8-10}$. We have made crystalline patches of PhoE porin embedded in a lipid bilayer and have used these to determine the structure of PhoE porin by electron crystallography to a resolution of 6 $\text{\AA}$. The basic structure consists of a trimer of elliptically shaped, cylindrical walls of β sheet. Each cylinder has an inner lining, formed by parts of the polypeptide, that defines the channel size. The structure provides a clue as to how deletions of segments of polypeptide, which are found in certain mutants, can result in an actual increase in the channel size.

We have reconstituted purified PhoE porin with phospholipids to form highly coherent, crystalline membrane patches that diffract to a resolution better than 3.0 $\text{\AA}$ (ref. 11). The crystalline membrane patches have a space group $P2_12_12$ ($a = 150 \text{ \AA}$ and $b = 129 \text{ \AA}$). A high-resolution projection map 12 and a three-dimensional (3-D) electron diffraction data set 11 have been obtained. The predominant features seen in the projection

FIG. 1 Amplitudes and phases along two lattice lines together with fitted curves. The smooth curves were obtained by weighted least-squares fitting as described by Agard¹⁸. The phase measurement is plotted with two different sizes of circles, a large circle (IQ 1 to 4)¹⁶ or a small circle (IQ 5 to 8), to indicate the reliability of the phase measurement. The amplitude measurements are derived from the earlier diffraction data. The fitted curves of phase and amplitude were used to generate the 3-D map.



map are high density, elliptical, ring-like structures which have been interpreted as the projected images of β sheet, seen edge-on. Each ring-like structure consists of 'beads', some of which have been interpreted as an end-on view of β strands. This interpretation is consistent with the diffraction data indicating that within PhoE porin a significant fraction of the β strands are oriented perpendicular to the membrane plane, whereas the predominant population of β strands are tilted by an average of 35° from the membrane normal¹¹. At the threefold symmetry axis of the trimer, there is a very-low-density zone that has been proposed as the location of lipopolysaccharide in the native membrane.

Electron diffraction patterns and images of PhoE porin crystals, embedded in trehalose and maintained at about -120°C , were recorded from samples tilted from 0 to about 60° , using

low-dose recording techniques. Both conventional flood beam illumination and spot-scan imaging techniques^{13,14} have been used. Images that were found to yield sharp, high-resolution spots by optical diffraction were selected, and digitized with a Perkin Elmer Model 1010M scanning densitometer. Images of tilted specimens were scanned with a $10\text{-}\mu\text{m}$ square aperture and a step size of $10\text{ }\mu\text{m}$ in arrays of 3,168 by 3,168 pixels. Twenty images at tilt angles ranging from 20° to about 60° were digitized and processed. The tilt axes and angles of images were determined as described by Shaw and Hills¹⁵. Computer processing of the digitized images followed the methods of Henderson *et al.*¹⁶, mainly using computer software from the MRC (Laboratory of Molecular Biology, Cambridge, UK). The processing included corrections of lattice distortions in the image, the effect of the contrast transfer function, and the defocus gradient of

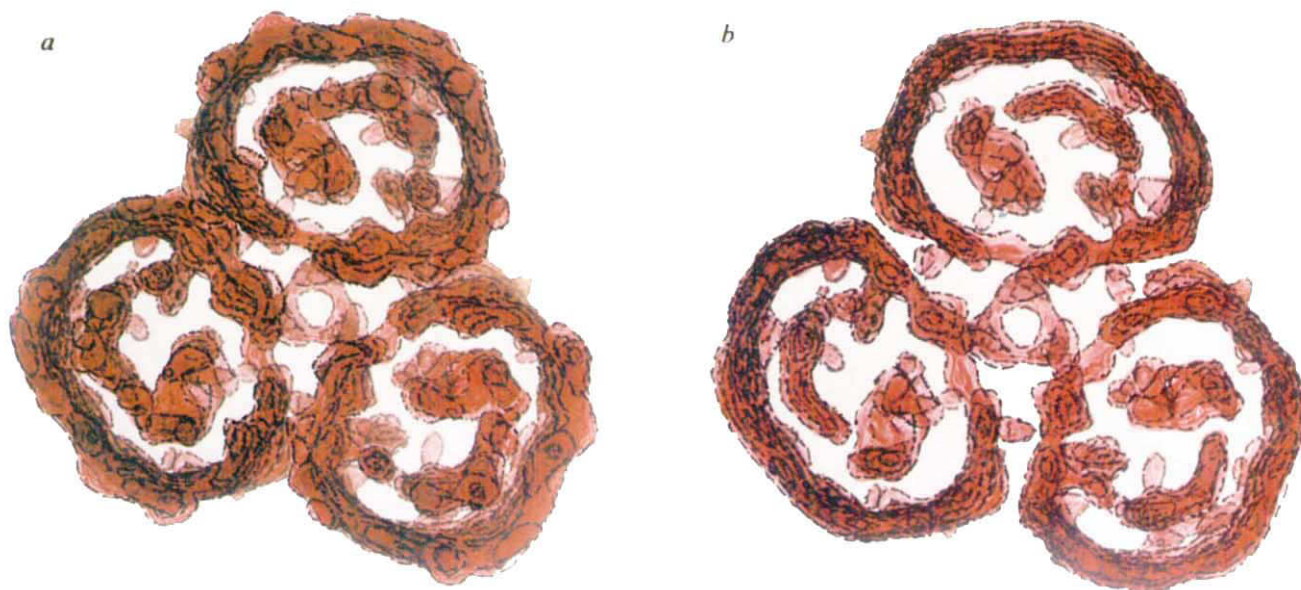


FIG. 2 Views of stacked serial sections obtained from the three-fold symmetrized 3-D map. Only the positive contour line (broken line) is shown here, representing the protein domains. The density in the map is contoured at an interval of $1/5$ of the maximum density, starting at a level of 20%. The protein domains are also highlighted in red to enhance the 3-D perspective. *a*, The stacked 3-D model of thickness $36\text{ }\text{\AA}$ viewed from the extracellular side of the structure. *b*, The bottom half of the structure. The maps show that the 3-D density consists of three elliptical, cylindrical walls that are arranged into a trimer. The elliptical cylinders have a major axis of $\sim 38\text{ }\text{\AA}$ and a minor axis of $\sim 28\text{ }\text{\AA}$ as measured from the centre of the wall. The cylinder wall extends to about $45\text{ }\text{\AA}$ in length; however, beyond about $35\text{ }\text{\AA}$,

the map is somewhat noisy. Each cylinder wall is clearly defined and there is no indication that the three channels of a trimer merge into a single pore. Within each cylinder, the substructure, also in our projection map¹², appears to extend along the entire length of the channel; the channel is not straight and is located off the elliptical cylinder axis. At the threefold symmetry axis, there is a low-density region that gradually decreases in size as it traverses from the extracellular to the periplasmic domain, reaching the smallest diameter about two-thirds of the way through the membrane. This region has previously been proposed to contain lipopolysaccharide on the extracellular side.

the tilted images¹⁶. The tilted images were brought to a common phase origin as described previously¹⁷. This was done by minimizing the difference between the phases of Fourier components of a given tilted image and the phases along the corresponding lattice lines determined previously. Initially, images with small tilt angles were merged with the projection map in which the phase values have been set to the nearest values of 0 or π , and then images with increasingly larger tilt angles were sequentially combined with the merged data. The phase errors resulting from such merging are less than 25° for each image to a resolution of 6 Å.

The 3-D map was reconstructed by Fourier transformation of hybrid structure factors created from the phase information derived from the images and more accurate amplitudes obtained from the 3-D electron diffraction data¹¹. The structure factors were derived from curves that were obtained from fitting the merged amplitude and phase data using the program LATLINE¹⁸. Fitting of curves to the data was performed with a weighted least-squares algorithm with a constraint that there is no electron density outside a 55 Å-thick envelope, surrounding the membrane. This width, somewhat wider than the lipid bilayer, was chosen to avoid the possibility of erroneous truncation of real structural density. The tilt axes and angles of all the images were further refined by minimizing the phase deviation between the experimental data and the smooth curves that were obtained from the initial curve fitting. These new tilt axes and angles were used to obtain the refined 3-D data set. Examples of plots of amplitude and phase and their fitted curves used in our 3-D reconstruction are shown in Fig. 1. In the current reconstruction, structure factors were truncated to a resolution of about 6 Å. As a result of not having data from specimens with tilt angles greater than 60 degrees, the effective resolution is about 6 Å parallel to the membrane plane and about 8 Å normal to the membrane.

The 3-D reconstruction of PhoE porin at 6 Å is shown in Fig. 2. The reconstruction shows several interesting features, which have previously been seen only in projection^{12,17}. The PhoE 3-D map confirms that three high-density, elliptical, cylindrical walls of β sheet are arranged as a trimer, with their cylinder axes almost parallel to each other and to the threefold axis, which is perpendicular to the membrane. The elliptical wall measures ~38 Å by ~28 Å along the major and minor axes, respectively. The length of the cylinder wall is ~45 Å, beyond ~35 Å, the map is somewhat noisy. Each cylinder wall is clearly defined, and there is no indication that the three channels of a trimer merge into a single pore, confirming our earlier interpretation of the negatively stained sample map¹⁷. Within each cylinder, additional density is found to be predominantly situated in contact with the part of the wall that is furthest away from the threefold symmetry axis, and this density appears to extend across the entire length of the channel. The open part of the channel therefore appears to be located off the central axis of the elliptical cylinder¹⁹, and the centre-to-centre distance between the apparent position of the aqueous channels is somewhat smaller than the distance between the centres of the trimeric, elliptical cylinders, which is about 45 Å. This is consistent with our negatively stained data showing that, within the trimer, the greatest centre-to-centre channel distance is about 35 Å. In our current map, it is rather difficult to define the channel shape clearly. The channel itself does not appear to run straight, which would be consistent with the negatively stained 3-D map, showing that the trimeric channels come close to each other as they traverse the membrane but they do not merge.

It is reasonable to believe that the additional density seen within each elliptical cylinder represents a region of the polypeptide that defines the channel size and is important in determining the channel selectivity. In fact, deletions of 6–15 residues in a specific region of the primary sequence, found in certain mutants, have been shown to greatly increase pore permeability^{20,21}. Point mutations in this region have also been found to

alter the pore permeability and selectivity. This suggests that this specific region of the sequence forms a part of the channel itself, and deletions in this segment would, therefore, result in significant increases in the channel size.

At the threefold symmetry axis of the trimer, there is a region of very low density that has been interpreted to be the site of lipopolysaccharide on the extracellular surface of the membrane in the native state. The presence of lipopolysaccharide at this region has also recently been reported by Hoenger *et al.*²². The low-density region is somewhat narrower in the periplasmic half of the membrane, indicating a reduced amount of lipid at the threefold symmetry axis on the periplasmic side of the membrane.

Recently, porin from *Rhodospseudomonas capsulata* has been crystallized, yielding X-ray diffraction intensities beyond 3.2 Å (ref. 23). With the use of X-ray crystallographic methods, a 3-D structural map of this trimeric porin has also been obtained to a resolution of 6 Å (ref. 24) and very recently its partial model structure at 3.0 Å resolution²⁵, with about three-quarters of the protein mass accounted for, has also been reported. The 6 Å map of this porin is different from that of PhoE porin, and does not show the well-defined cylindrical β -sheet wall which is seen in our 3-D map of PhoE porin or in the projection maps of PhoE and OmpF porins^{17,26}. The 3-D map of *R. capsulata* porin shows two possible channels per porin monomer: one having a diameter of about 6–10 Å and the other somewhat narrower, with a diameter of about 5 Å. The centre-to-centre separation of the larger *R. capsulata* channels is roughly 44 Å, which is somewhat greater than that of PhoE porin as derived from either our current 6 Å map or our previous negatively stained 3-D map. Interestingly, the structure of *R. capsulata* porin at 3.0 Å resolution does show some similarity to that of PhoE porin at 6 Å resolution. Both porins consist of elliptical, cylindrical walls of β sheet that are arranged into a trimer. There are also many differences. The cylindrical walls forming the monomer-monomer contact region of the trimer of *R. capsulata* is significantly shorter than the overall length of the whole cylinder. Therefore, the trimeric channels form a common chamber at one side of the channel. In contrast, each cylinder wall of PhoE porin is clearly defined in the region of monomer-monomer contact, and there is no indication that the three channels of a trimer merge into a single pore. Within each cylinder, the substructure of PhoE porin appears to extend across the entire length of the channel although that of *R. capsulata* porin is confined around the central region of the channel. Another significant difference between these two porins can be found in their oligomeric states. In contrast to other porins that normally form very stable trimers, *R. capsulata* porins have been found in monomeric form even under mild purification conditions. The question of the similarity between the architecture of these two porins will have to await the availability of their respective detailed molecular structures. □

Received 4 December 1990; accepted 4 January 1991.

1. Nikaido, H. & Vaara, M. *Microbiol. Rev.* **49**, 1–32 (1985).
2. Lugtenberg, B. & Van Alphen, L. *Biochim. biophys. Acta* **737**, 51–115 (1983).
3. Schindler, H. & Rosenbusch, J. P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3751–3755 (1978).
4. Dargent, B., Hofmann, W., Pattus, F. & Rosenbusch, J. P. *EMBO J.* **5**, 773–778 (1986).
5. Xu, G., Shi, B., McGroarty, E. J. & Tien, H. T. *Biochim. biophys. Acta* **862**, 57–64 (1986).
6. Mauro, A., Blake, M. & Labarca, P. *Proc. natn. Acad. Sci. U.S.A.* **85**, 1071–1075 (1988).
7. Kortelant, J. & Tommassen, J. & Lugtenberg, B. *Biochim. biophys. Acta* **690**, 282–289 (1982).
8. Rosenbusch, J. P. *J. biol. Chem.* **249**, 8019–8029 (1974).
9. Kleffel, B., Garavito, R. M., Baumeister, W. & Rosenbusch, J. P. *EMBO J.* **4**, 1589–1592 (1985).
10. Tommassen, J. *Membrane Biogenesis* NATO ASI Series Vol. 16 (ed. Op den Kamp J.A.F.) 351–373 (NATO 1988).
11. Walian, P. J. & Jap, B. K. *J. molec. Biol.* **215**, 429–438 (1990).
12. Jap, B. K., Downing, K. H. & Walian, P. J. *J. struct. Biol.* **103**, 57–63 (1990).
13. Downing, K. H. & Glaeser, R. M. *Ultramicroscopy* **20**, 269–278 (1986).
14. Bullough, P. & Henderson, R. *Ultramicroscopy* **21**, 223–230 (1987).
15. Shaw, P. J. & Hills, G. *J. Microsc.* **12**, 279–282 (1981).
16. Henderson, R. *et al. J. molec. Biol.* **213**, 899–929 (1990).
17. Jap, B. K. *J. molec. Biol.* **205**, 407–419 (1989).
18. Agard, D. A. *J. molec. Biol.* **167**, 849–852 (1983).
19. Jap, B. K. & Walian, P. J. *Q. Rev. Biophys.* **23**, 4, 367–403 (1990).
20. Misra, R. & Benson, S. A. *J. Bact.* **170**, 3611–3617 (1988).

21. Benson, S. A., Occhi, J. L. & Sampson, B. A. *J. molec. Biol.* **203**, 961-970 (1988).
22. Hoenger, A., Gross, H., Aebi, U. & Engel, A. *J. struct. Biol.* **103**, 185-195 (1990).
23. Nestel, U. *et al.* *FEBS Lett.* **242**, 405-408 (1989).
24. Weiss, M. S. *et al.* *FEBS Lett.* **256**, 143-146 (1989).
25. Weiss, M. S., Wacker, T., Weckesser, J., Welte, W. & Schulz, G. E. *FEBS Lett.* **267**, 268-272 (1990).
26. Sass, H. J. *et al.* *J. molec. Biol.* **209**, 171-175 (1989).

ACKNOWLEDGEMENTS. We thank K. H. Downing and R. M. Glaeser for helpful discussions. This work is supported by the Office of Health and Environmental Research, US Department of Energy and the NIH.

A haemoprotein with kinase activity encoded by the oxygen sensor of *Rhizobium meliloti*

Marie A. Gilles-Gonzalez, Gary S. Ditta & Donald R. Helinski

Department of Biology and Center for Molecular Genetics, 0634, University of California, San Diego, La Jolla, California 92093, USA

THE expression of the nitrogen-fixation genes of *Rhizobium meliloti* is controlled by oxygen^{1,2}. These genes are induced when the free oxygen concentration is reduced to microaerobic levels. Two regulator proteins, FixL and FixJ, initiate the oxygen-response cascade, and the genes that encode them have been cloned^{2,3}. The *fixL* product seems to be a transmembrane sensor that modulates the activity of the *fixJ* product, a cytoplasmic regulator³. FixL and FixJ are homologous to a family of bacterial two-component regulators⁴, for which the mode of signal transduction is phosphorylation (reviewed in refs 5-9). We report here the purification of both FixJ and a soluble truncated FixL (FixL*), overproduced from a single plasmid construct. FixL* catalyses its own phosphorylation and the transfer of the γ -phosphate of ATP to Fix J. The resulting FixJ-phosphate linkage is sensitive to base, as are the aspartyl phosphates of homologous systems. Visible spectra of purified FixL* show that it is an oxygen-binding haemoprotein. We propose that FixL senses oxygen through its haem moiety and transduces this signal by controlling the phosphorylation of FixJ.

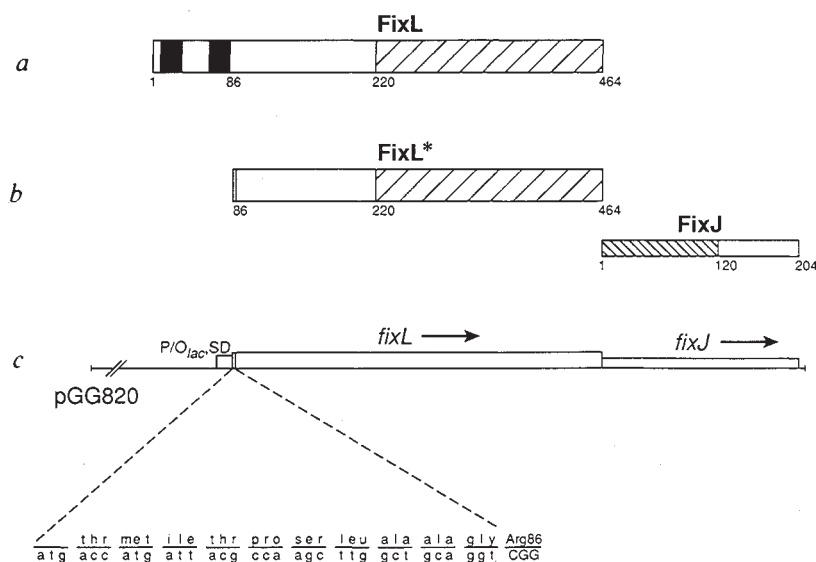
Two long stretches of hydrophobic amino acids near the N terminus of FixL were deleted, to facilitate purification (Fig. 1a and b). This truncated form of FixL, designated FixL*, was

coexpressed in *Escherichia coli* with FixJ, because FixL and FixJ might stabilize each other's activity. This was achieved by constructing a plasmid in which a 5' truncated *fixLJ* operon was placed under the control of a *lac* promoter-operator (Fig. 1c). Two new proteins having relative molecular masses (M_r s) expected for either FixL* (M_r 42,803) or FixJ (M_r 22,849) were revealed on electrophoresis of whole-cell lysates from isopropyl- β -D-thiogalactoside (IPTG) induced, but not from non-induced, cells carrying the cloned genes. The presence of FixL* was confirmed by western immunoblot analysis using an antibody specific for the newly induced protein of M_r ~43,000 (43K). FixJ was the only 23K phosphoprotein that could be detected and was present exclusively in induced cells that carried the cloned genes. It also reacted to FixJ-specific antibody.

Both FixL* and FixJ partitioned to the soluble fraction of the sonicated cells (Fig. 2, lane a). They were purified more than 10-fold, as measured by protein assays using ion-exchange chromatography (Fig. 2, lane b). A peak of orange-red fractions was apparent. Gel electrophoresis and phosphorylation assays showed that these fractions contain both proteins. (On subsequent purification, the colour stayed with FixL*) FixJ (>90% pure; about 10 mg from 20 g cells) was resolved from FixL* by gel filtration, indicating that they were not stably complexed (Fig. 2, lanes c and d). The native M_r s estimated from gel filtration were 50K for FixJ and 93K for FixL*, close to the values expected for dimeric FixJ (46K) and dimeric FixL* (86K). Orange-red FixL* of >95% purity (about 11 mg from 20 g cells) was obtained after further purification on MonoQ (Fig. 2, lane e).

Purified FixL* was reacted with [γ -³²P]ATP and examined by SDS-PAGE (Fig. 3a) and autoradiography (Fig. 3b). Radiolabel was detected in FixL*, indicating that FixL* can covalently attach to the γ phosphate of ATP (Fig. 3a and b, lane a). By contrast, no radiolabel was found in FixJ alone incubated under the same conditions (Fig. 3a and b, lane b). When the two proteins were incubated together with ATP, both FixL* and FixJ were labelled (Fig. 3a and b, lane c), indicating that FixL* is a kinase that can catalyse both its own phosphorylation and phosphorylation of Fix J. When the phosphorylated proteins were exposed to 0.5 M NaOH at 55 °C for 5 min, all the radiolabelled phosphate was hydrolysed from phospho-FixJ, whereas phospho-FixL was largely unreacted (Fig. 3a and b, lane d). Exposure of phospho-FixJ to 55 °C for 5 min under neutral conditions did not result in hydrolysis of the phosphate (data

FIG. 1 Recombinant *fixLJ* region and its protein products. a, The structure of full-length FixL is schematically represented. ■, Hydrophobic stretches of amino acids. ▨, Regions of homology to other two-component systems. The numbers refer to amino-acid positions from David *et al.*³. b, Structures of FixL* and FixJ purified in this study. Positions of relevant amino acids are indicated. c, Recombinant genes that encode FixL* and FixJ. The DNA and amino-acid sequences of the *lacZ*-*fixL* junction are also shown. The genes are on a high-copy-number plasmid, pGG820, derived from pUC9. □, The *lac* promoter, operator and ribosome-binding site (*lac* P/O, RBS); □, sequences encoding FixL* and FixJ. METHODS. Plasmid pGG820 was obtained in three steps. The initial plasmid for the constructions was pMW2 (M. Weinstein, unpublished data). Plasmid pMW2 consists of a 3.15-kilobase (kb) *Bam*HI fragment of *R. meliloti* DNA, containing the *fixLJ* operon, cloned into the *Bam*HI site of pUC9. First, plasmid pGG8, which contains only pUC9 sequences and the 5' region of *fixL* was constructed. A 1.88-kb *Eco*RI fragment containing the 3' region of the *fixLJ* operon was excised from pMW2, and pGG8 was produced by religating the vector fragment with T4 DNA ligase. In pGG82, the 86th codon of *fixL* is fused to the pUC9 polylinker region in-frame with first six *lacZ* codons. This was accomplished by excising a 0.63-kb *Acc*I-*Rsr*II region from pGG8, treating the vector fragment with Klenow and dNTPs to generate blunt DNA ends, then religating with T4 DNA ligase. Finally, the 3' region of the *fixLJ* operon was reconstructed in pGG820



by cutting pGG82 with *Eco*RI and ligating it to the 1.88-kb *Eco*RI fragment from pMW2. The relevant regions of pGG820 are shown in Fig. 1c.

not shown). These results are consistent with the presence of an acyl phosphate linkage in phospho-FixJ and with the properties of the phosphorylated regulator proteins of other prokaryotic regulatory systems¹⁰⁻¹².

The visible spectrum of a solution of purified FixL* from 350

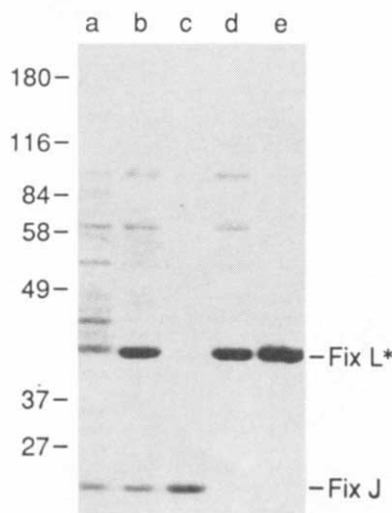


FIG. 2 Purification of FixL* and FixJ. Protein from each purification step was examined by electrophoresis on SDS-polyacrylamide gels (11%). Each lane contains 5 μ g. The samples were: high-spin supernatant (lane a); fractions pooled after DEAE-Sephacel (lane b); FixJ fractions pooled after Sephacryl S-200 (lane c); FixL* fractions pooled after Sephacryl S-200 (lane d); FixL* fractions pooled after FPLC on a MonoQ column (lane e). M_r markers ($M_r \times 10^{-3}$; left) and the bands corresponding to FixL* and FixJ are indicated. **METHODS.** The *E. coli* host strain was TG1 [*supE hsd Δ 5 thi Δ (lac proAB) F'(traD36 proAB⁺ lacI^q lacZ Δ M15)]¹⁴. This strain was transformed with the *fixLJ* plasmid pGG820 and the *nifA* promoter test plasmid pCHK57⁷. Strain TG1(pCHK57, pGG820) was maintained in M9 medium supplemented with 10 mM glucose, 1 μ g ml⁻¹ thiamine, 250 μ g ml⁻¹ penicillin G, and 15 μ g ml⁻¹ tetracycline. Before induction, a concentrated inoculum was prepared from a logarithmic culture grown in the above medium at 37 °C. LB, containing 250 μ g ml⁻¹ penicillin G and 15 μ g ml⁻¹ tetracycline, was inoculated to a starting 600 nm absorbance of 0.05. When the absorbance at 600 nm of the culture reached 0.4–0.5, IPTG was added to a final concentration of 1 mM, and the cells were collected after 5 h at 37 °C. A cell pellet obtained from 5 litres culture (20 g wet cells) was washed then resuspended in 80 ml buffer A (20 mM Tris-HCl, pH 7.8, 125 mM NaCl, 5% (v/v) glycerol, 10 mM β -mercaptoethanol, and 1 mM PMSF). The cells were lysed by 3-min sonication. Unbroken cells were removed by centrifugation at 3,000 r.p.m. (Sorvall rotor SS34) for 10 min. The crude extracts were then spun at 45,000 r.p.m. for 30 min (Ti60 rotor). The supernatant was directly applied to a DEAE-Sephacel column (2.5 $\times$ 12 cm) pre-equilibrated with buffer A, and the column was washed with 150 ml of the same buffer. Proteins were then eluted with 400 ml of a linear NaCl gradient from 125 to 400 mM, at a flow rate of 30 ml h⁻¹. The fractions containing FixL* and FixJ eluted at around 200 mM NaCl. The pooled fractions (50 ml) were concentrated to a volume of 1 ml by precipitation from 40% saturated ammonium sulphate and applied onto a Sephacryl S-200 column (2 $\times$ 125 cm) equilibrated with 20 mM Tris-HCl, pH 7.8, 150 mM NaCl, 5% glycerol and 10 mM β -mercaptoethanol. Elution was performed with the same buffer at a flow rate of 3 ml h⁻¹. The fractions having pure FixJ were combined and stored at -70 °C. Purified FixJ reacted positively with FixJ-specific antibody prepared by A. Lois. The fractions containing FixL* were further purified on a FPLC MonoQ HR5/5 1-ml column equilibrated with buffer B (20 mM Tris-HCl, pH 9.0, 125 mM NaCl, 5% glycerol and 10 mM β -mercaptoethanol). Protein was loaded onto the column in buffer B and eluted with a 25-ml linear 125–750 mM NaCl gradient at a flow rate of 30 ml h⁻¹. The fractions that eluted at 480 mM NaCl contained pure FixL*. Purification procedures were done at 4 °C. All chromatographic media were purchased from Pharmacia. Protein was measured by the Bradford method (Bio-Rad), using BSA as the standard. Initially, both FixL* and FixJ were detected by phosphorylation assays. Later, FixL* was monitored by absorption measurements near its absorption maximum (420 nm). The Sephacryl S-200 column was calibrated with a mixture of the M_r standards β -amylase (200K), alcohol dehydrogenase (150K), BSA (66K), carbonic anhydrase (29K) and cytochrome *c* (12.4K).*

to 650 nm is that of a haemoprotein (Fig. 4). The absorption maxima of aerobic FixL*, at 576.5 (α), 542.5 (β) and 417 nm (γ , or Soret) and their relative heights are characteristic of haemoproteins. Oxyhaemoglobin, for example, has peaks of 576 (α), 541 (β) and 415 (γ) (Fig. 4a). Deoxygenation of the FixL*

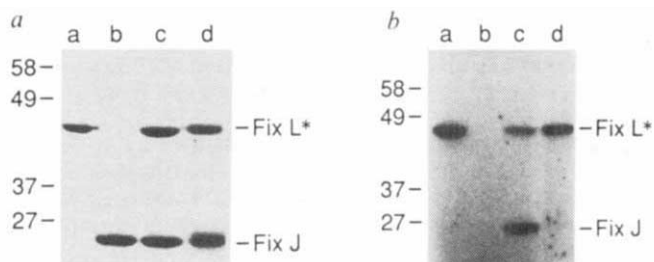


FIG. 3 Phosphorylation of FixL* and FixJ with [γ -³²P]ATP. Each purified protein present (1.9 μ M FixL*, 3.6 μ M FixJ) was incubated for 5 min in standard phosphorylation buffer in air. FixL* alone, lane a; FixJ alone, lane b; FixL* plus FixJ, lane c; FixL* plus FixJ treated 5 min at 55 °C in 0.5 M NaOH after phosphorylation, lane d. Products were analysed by Coomassie brilliant blue staining (a) or autoradiography (b) of SDS-polyacrylamide gels of the reaction mixtures after electrophoresis. The Coomassie-stained lanes had three times more sample. M_r markers ($M_r \times 10^{-3}$; left) and the positions of FixL* and FixJ are shown.

METHODS. Protein (2–7 μ M) in 12.5 μ l 60 mM Tris-HCl, pH 8.0, 60 mM KCl, 5% (v/v) glycerol and 1.2 mM dithiothreitol was equilibrated for 5 min at 23 °C. Phosphorylation was initiated by adding 2.5 μ l of a solution of 2.4 mM unlabelled ATP, 2.4 mM MgCl₂ and 10 μ Ci [γ -³²P]ATP (Amersham). After 5 min of incubation at room temperature, the reaction was stopped with 5 μ l SDS buffer (125 mM Tris-HCl, pH 6.8, 8 mM EDTA, 4% (w/v) SDS, 8% (v/v) β -mercaptoethanol, 20% (v/v) glycerol and 0.02% (w/v) bromophenol blue). To study the stability of the phosphate linked to FixJ, the reaction products were treated with 5 μ l 2 M NaOH, 8 mM EDTA for 5 min at 55 °C. The mixture was then neutralized with HCl, and SDS buffer was added. The samples were electrophoresed on an SDS-polyacrylamide gel (10–11%). The gel was dried in a vacuum, and the labelled proteins were detected by autoradiography.

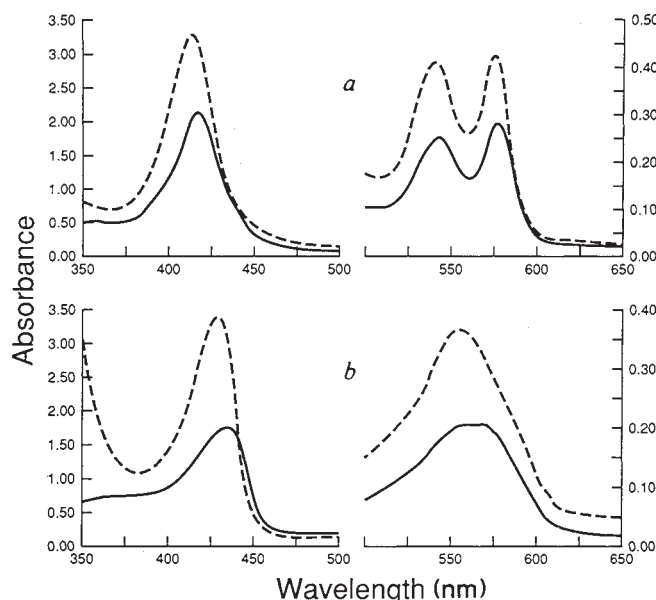


FIG. 4 Spectra of oxygenated and deoxygenated FixL*. Aerobic FixL* (13.4 μ M monomer) in 20 mM Tris-HCl, pH 8.5, 150 mM NaCl, 5% (v/v) glycerol and 10 mM β -mercaptoethanol (a, solid curve); identical FixL* solution after it was made anaerobic (b, solid curve); oxyhaemoglobin (a, dashed curve); deoxyhaemoglobin (b, dashed curve). Wavelength scans were collected on a Uvicon 930 spectrophotometer (Kontron Instruments) in 1-cm path length quartz cells.

haem *in vacuo* resulted in absorption shifts very similar to those observed for deoxyhaemoglobin (Fig. 4b). The α and β bands are replaced by a single peak at 567.5 nm, and the Soret band is shifted to 433.5 nm. The corresponding deoxyhaemoglobin peaks were at 555 and 430 nm (Fig. 4b). These changes were completely reversible, demonstrating that FixL* retains its ability to respond to oxygen. The extinction coefficient of monomeric oxy-FixL* at 417 nm is roughly $1.6 \times 10^5 \text{ M}^{-1} \text{ cm}^{-1}$, on the basis of protein assays and absorbance at 417 nm. Pyridine haemochrome assay suggests that FixL* contains one mole of haem per monomer.

FixL* is a novel haemoprotein (Fig. 4) with kinase activity (Fig. 3). We propose that FixL has at least three distinct regions (Fig. 1). The homologous region (residues 220–464) houses those functions that are common to the two-component sensors, such as ATP binding, partner recognition and phosphorylation (reviewed in refs 5–9). The hydrophobic region (residues 1–85) probably anchors FixL to the cell membrane. These residues, absent in FixL*, are unnecessary for haem binding or phosphorylation. They may stabilize the protein *in vivo* or protect the haem from oxidation. Membrane attachment may also enhance or modify the functions of FixL, or FixL may communicate with other membrane proteins. An extracytoplasmic loop could lie in the deleted hydrophobic region of FixL*. The removal of the extracytoplasmic loop from VirA, in the agrobacterial virulence system, did not eliminate the *in vivo* response to acetosyringone, indicating that the site of acetosyringone-sensing could be intramembranous or cytoplasmic¹³. The remaining region of FixL (residues 86–219) is likely to contain the haem attachment site, that is, the sensing portion of the

molecule. This region is most probably on the cytoplasmic side of the membrane, because no hydrophobic stretches separate it from the enzymatic domain. Histidines are expected to have an important role in the function of this molecule, because they are commonly involved in haem coordination and are the sites of phosphorylation in several proteins homologous to FixL^{10,15,16}. That FixL is a haemoprotein with kinase activity raises the question of whether the oxygenation state of the haem moiety affects its kinase activity. Experiments are in progress to answer this question and to provide evidence for a linkage between the phosphorylation events involving FixL and FixJ and signal transduction in nitrogen fixation. □

Received 12 November 1990; accepted 22 January 1991.

1. Ditta, G., Virts, E., Palomares, A. & Kim, C.-H. *J. Bact.* **169**, 3217–3223 (1987).
2. Virts, E. L., Stanfield, S. W., Helinski, D. R. & Ditta, G. S. *Proc. natn. Acad. Sci. USA* **85**, 3062–3065 (1988).
3. David, M. *et al. Cell* **54**, 671–683 (1988).
4. Nixon, B. T., Ronson, C. W. & Ausubel, F. M. *Proc. natn. Acad. Sci. USA* **83**, 7850–7854 (1986).
5. Stock, A. M. *et al. Cold Spring Harb. Symp. quant. Biol.* **53**, 49–57 (1989).
6. Stock, J. B., Ninfa, A. J. & Stock, A. M. *Microbiol. Rev.* **53**, 450–490 (1989).
7. Stock, J. B., Stock, A. M. & Mottonen, J. M. *Nature* **344**, 395–400 (1990).
8. Albright, L. M., Huala, E. & Ausubel, F. M. *A. Rev. Genet.* **23**, 311–336 (1989).
9. Gross, R., Arico, B. & Rappouli, R. *Molec. Microbiol.* **3**, 1661–1667 (1989).
10. Weiss, V. & Magasanik, B. *Proc. natn. Acad. Sci. USA* **85**, 8919–8923 (1988).
11. Forst, S., Delgado, J. & Inouye, M. *Proc. natn. Acad. Sci. USA* **86**, 6052–6056 (1989).
12. Bourret, R. B., Hess, J. F. & Simon, M. L. *Proc. natn. Acad. Sci. USA* **87**, 41–45 (1990).
13. Melchers, L. S. *et al. EMBO J.* **8**, 1919–1925 (1989).
14. Gibson, T. J. thesis, Cambridge Univ., UK (1984).
15. Hess, J. F., Bourret, R. B. & Simon, M. L. *Nature* **336**, 139–143 (1988).
16. Jin, S. *et al. J. Bact.* **172**, 525–530 (1990).

ACKNOWLEDGEMENTS. We thank P. Saltman and B. Dyke for discussions, comments on the manuscript and assistance with the spectroscopy; P. Geidushek for comments on the manuscript; and M. Adams and G. Anders for help with immunizations. This work was supported by the NSF. M.G.-G. is supported by a Fellowship from the NIH.

Single-site enzymatic cleavage of yeast genomic DNA mediated by triple helix formation

Scott A. Strobel & Peter B. Dervan*

Division of Chemistry and Chemical Engineering,
California Institute of Technology, Pasadena,
California 91125, USA

PHYSICAL mapping of chromosomes would be facilitated by methods of breaking large DNA into manageable fragments, or cutting uniquely at genetic markers of interest. Key issues in the design of sequence-specific DNA cleaving reagents are the specificity of binding, the generalizability of the recognition motif, and the cleavage yield. Oligonucleotide-directed triple helix formation is a generalizable motif for specific binding to sequences longer than 12 base pairs within DNA of high complexity^{1–3}. Studies with plasmid DNA show that triple helix formation can limit the operational specificity of restriction enzymes to endonuclease recognition sequences that overlap oligonucleotide-binding sites^{4,5}. Triple helix formation, followed by methylase protection, triple helix-disruption, and restriction endonuclease digestion produces near quantitative cleavage at the single overlapping triple helix–endonuclease site^{4,5}. As a demonstration that this technique may be applicable to the orchestrated cleavage of large genomic DNA, we report the near quantitative single-site enzymatic cleavage of the *Saccharomyces cerevisiae* genome mediated by triple helix formation. The 340-kilobase yeast chromosome III was cut uniquely at an overlapping homopurine–*EcoRI* target site 27 base pairs long to produce two expected cleavage products of 110 and 230 kilobases. No cleavage of any other chromosome was detected. The potential generalizability of this technique, which is capable of near quantitative cleavage at a single site in at least

14 megabase pairs of DNA, could enable selected regions of chromosomal DNA to be isolated without extensive screening of genomic libraries.

Chemical reagents and biological methods that infrequently cleave double helical DNA are being developed for genomic mapping^{1–12}. Recently, Szybalski and coworkers reported an 'Achilles heel cleavage' technique that limits restriction enzyme cleavage to an overlapping *lac* repressor–endonuclease site^{11,12}. Protein-mediated Achilles heel cleavage may, however, not be readily generalized to unique cleavage at selected genetic markers owing to the paucity of applicable DNA-binding proteins relative to the number of sites in megabase genomic DNA. To achieve orchestrated cleavage at selected genetic markers in human DNA by this approach may require the protein-binding sequence to be artificially inserted into the chromosome at precise locations. A more general method for recognition of endogenous DNA sequences might be a chemical approach based on oligonucleotide-directed triple helix formation¹. Pyrimidine oligonucleotides bind specifically to purine sequences in duplex DNA to form a local triple-helix structure^{1–5,13,14}. The generalizable code for triple-helix specificity is derived from thymine (T) binding to adenine–thymine base pairs (T·AT base triplet)¹⁵ and N3 protonated cytosine (C+) binding to guanine–cytosine base pairs (C+GC base triplet)^{16,17}. Higher affinity oligonucleotides can be obtained by substituting 5-bromouracil for thymine and 5-methylcytosine for cytosine¹⁸. Moreover, the number of potential target sequences amenable to recognition by the triple helix motif can be extended to some mixed purine–pyrimidine sequences^{19,20}. By replacing the *lac* repressor protein in the Achilles heel technique with triple-helix-mediated protection, one combines the strengths of the two approaches; a general chemical method for recognition of DNA sequences in the range of 15–20 base pairs (bp) with the cleavage efficiency of restriction enzymes (Fig. 1).

To demonstrate single-site enzymatic cleavage of the yeast genome by this technique, a sequence containing an overlapping 24-bp purine tract and 6-bp *EcoRI* site was inserted proximal to the *LEU2* (ref. 21) gene on the short arm of chromosome

* To whom correspondence should be addressed.

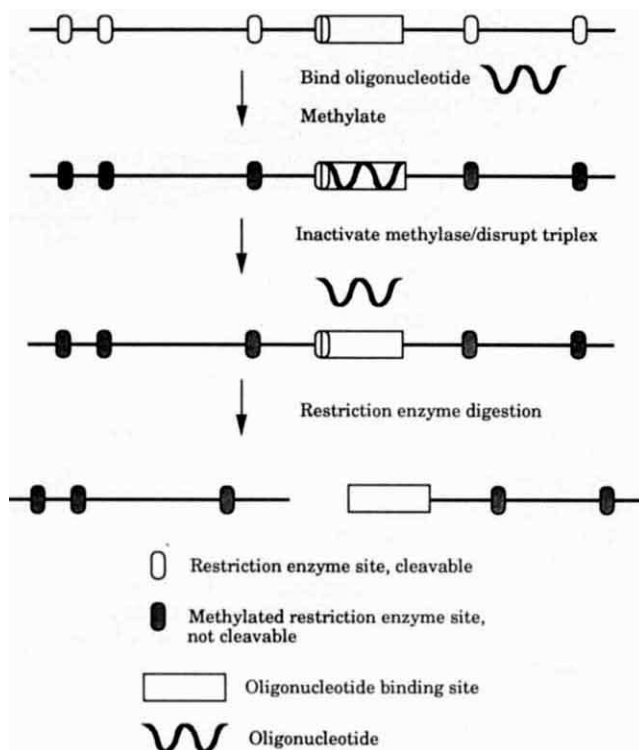


FIG. 1. General scheme for single-site enzymatic cleavage of genomic DNA by oligonucleotide-directed triple helix formation. Chromosomal DNA is equilibrated with an oligonucleotide in a methylase compatible buffer containing polyclonal. *EcoRI* methylase which methylates the central adenines of the sequence 5'-GAATTC-3' and renders the sequence resistant to cleavage by *EcoRI* restriction endonuclease, is added and allowed to proceed to completion. The methylase is inactivated and the triple helix is disrupted at 55 °C in a high-pH buffer containing detergent. After washing extensively, the chromosomal DNA is re-equilibrated in restriction enzyme buffer and cut to completion with *EcoRI* restriction endonuclease. The cleavage products are separated by pulsed-field gel electrophoresis and efficiencies quantitated by Southern blotting.

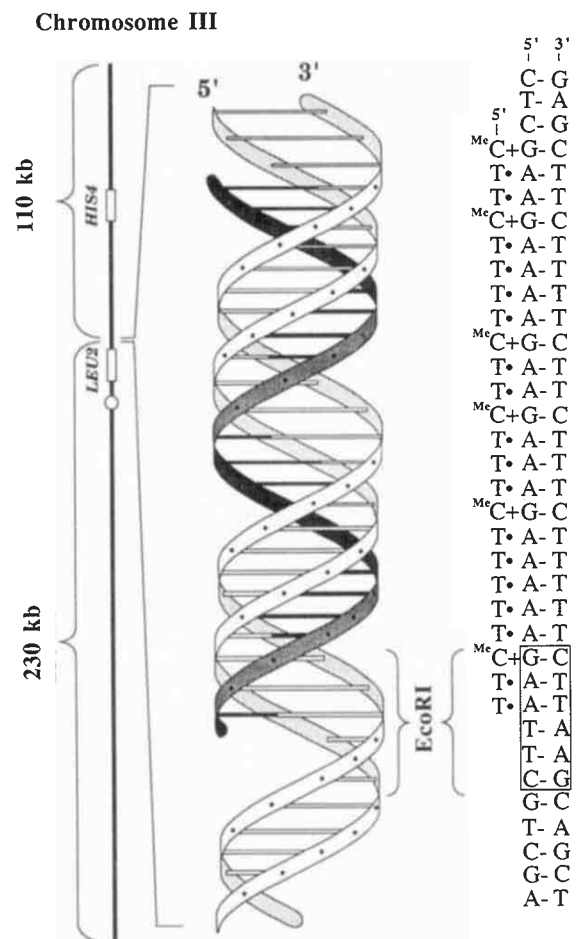


FIG. 2 Left, genetic map of *S. cerevisiae* chromosome III. The locations of the *HIS4* and *LEU2* loci (boxes), the centromere (circle), and the triple helix-*EcoRI* target site are indicated. The expected sizes of the cleavage products are shown. Right, schematic diagram of the triple helix complex overlapping the *EcoRI* restriction-methylation site. The pyrimidine oligonucleotide is bound in the major groove parallel to the purine strand of the DNA duplex, and covers half the *EcoRI* site.

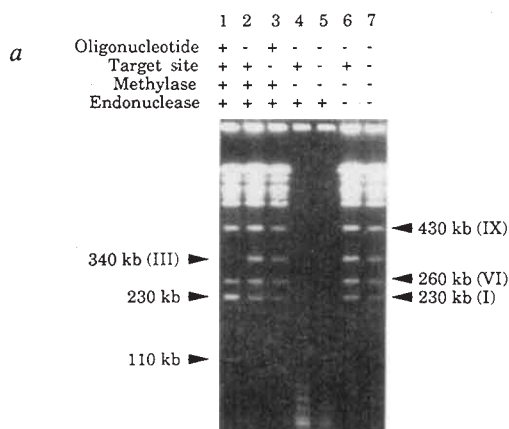
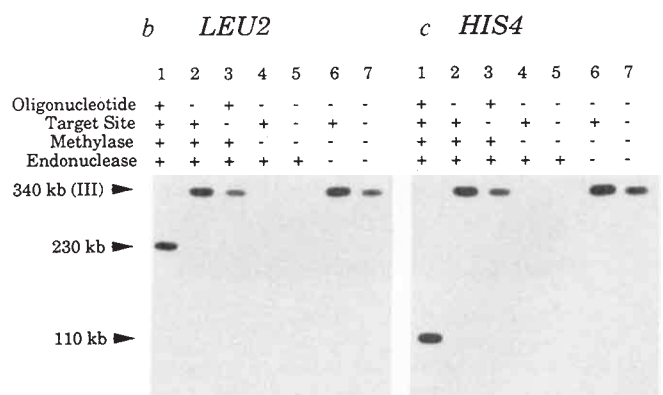


FIG. 3 Single-site enzymatic cleavage of yeast genomic DNA with reagents indicated above figure. 1, Yeast chromosomal DNA embedded in low-melting-point agarose ($\sim 50 \mu\text{l}$) was washed twice in 1.0 ml of 100 mM NaCl, 100 mM Tris-HCl, 10 mM EDTA, 2 mM spermine, pH 7.6 for 10 min, decanted, and overlaid with 150 μl of the same buffer. Oligonucleotide was added to 1 μM final concentration and incubated 8 h at 22 $^{\circ}\text{C}$. 2, Bovine serum albumin (100 $\mu\text{g ml}^{-1}$), *S*-adenosylmethionine (160 μM) and *EcoRI* methylase (80 units) were added to the overlay and incubated with shaking for 3 h. 3, The triple helix was destabilized and the methylase was simultaneously inactivated at 55 $^{\circ}\text{C}$ in 1.0 ml of 1% lauryl sarcosyl, 100 mM Tris-HCl pH 9.5, 10 mM EDTA for 30 min. The oligonucleotide and detergent were then removed with 4 $\times$ 10 min washes (1.0 ml) with 10 mM Tris-HCl pH 9.5, 10 mM



EDTA. 4, The agarose plug was washed twice with 1.0 ml 2×potassium glutamate restriction enzyme buffer (200 mM potassium glutamate, 50 mM Tris-acetate, pH 7.6, 20 mM magnesium acetate, 100 µg ml⁻¹ BSA, 1 mM 2-mercaptoethanol), decanted, overlaid with 150 µl of buffer and the DNA digested to completion with 20 U *Eco*RI restriction endonuclease for 6 h at 22 °C. After digestion the enzyme was heat-inactivated at 55 °C for 10 min and the DNA loaded onto a 1% agarose, 0.5×TBE pulsed-field gel at 14 °C, 200 V. Switch times were ramped from 10–40 s for 18 h followed by ramping from 60–90 s for 6 h. Products were visualized by ethidium bromide staining (a) and Southern blotting with *LEU2* (b) and *HIS4* (c) chromosome markers. Cleavage efficiencies were measured using storage phospho imaging plates with a Molecular Dynamics 400S Phosphorager.

III (Fig. 2) (ref. 3). Oligonucleotides designed to form triple helix complexes that overlap half of the *EcoRI* recognition site were synthesized with CT, ^{Me}CT or ^{Me}CBrU nucleotides¹⁸. The genetic map of yeast chromosome III (refs 22, 23) and affinity cleaving data³ indicate that cleavage at the target site should produce two fragments 110±10 and 230±10 kilobases (kb) in size (Fig. 2).

Resolution of total yeast chromosomal DNA by pulsed-field gel electrophoresis^{24,25} revealed that chromosome III was cut exclusively at the target site when a ^{Me}CT oligonucleotide was used for triple helix formation at pH 7.6 (Fig. 3, lane 1). No cleavage was detected on any other chromosomes under these conditions nor was cleavage observed in the absence of oligonucleotide (lane 2) or in a yeast strain lacking the target sequence (lane 3). The expected 110-kb product was visualized with ethidium bromide staining and confirmed by Southern blotting with a *HIS4* (ref. 26) marker (Fig. 3c). The 230-kb product comigrated with chromosome I, but was detected by Southern blotting with a *LEU2* (ref. 24) marker (Fig. 3b). The cleavage efficiency was 94±2%. Similar efficiencies were seen with a CT oligonucleotide up to pH 7.4 and ^{Me}CT and ^{Me}CBrU oligonucleotides past pH 7.8 (Fig. 4). The cleavage efficiency with all oligonucleotides was gradually reduced with longer methylation times, suggesting that the oligonucleotide dissociation rate might be the limiting factor for efficiency in this system²⁷.

The specificity of triple helix formation has been shown to be pH-dependent¹⁻³. By lowering the pH, sequences of near but imperfect similarity can be bound and cleaved¹⁻³. In agreement with this observation, secondary cleavage sites were revealed as a function of oligonucleotide composition and pH (Fig. 4). Cleavage at the secondary site was preferentially reduced by

longer methylation times, suggesting differential oligonucleotide dissociation rates between the primary and secondary target sites. Cleavage at all secondary sites could be eliminated at a threshold pH for each oligonucleotide.

Whereas affinity cleaving using oligonucleotide-EDTA-Fe identifies all sites of oligonucleotide binding¹⁻³, triple-helix-mediated endonuclease cleavage exposes only those sequences that also partially overlap a restriction site^{4,5}. The sequence requirement of a methylation-restriction site increases the cleavage specificity but reduces the number of available sites. To partially overcome this limitation, other commercially available methylation and restriction enzyme pairs with homopurine half sites were tested on plasmid DNA. In addition to *TaqI* (ref. 4) and *EcoRI* (ref. 5), single-site protection of plasmid DNA was possible with *MspI*, *HpaII* and *AluI* methylases. *BamHI*, *HaeIII* and *dam* methylases could potentially be used, but remain untested.

The generalizability of triple helix-mediated enzymatic cleavage affords high specificity that can, in principle, be customized to unique genetic markers without artificial insertion of a target sequence. The use of degenerate oligonucleotides in this technique to rapidly screen genetic markers for overlapping triple-helix methylation-restriction sites could make it possible to cut chromosomal DNA uniquely and efficiently at endogenous sites with minimal sequence information (S.A.S. and P.B.D., unpublished observations). □

Received 5 October; accepted 20 December 1990.

1. Moser, H. E. & Dervan, P. B. *Science* **238**, 645-650 (1987).
2. Strobel, S. A., Moser, H. E. & Dervan, P. B. *J. Am. chem. Soc.* **110**, 7927-7929 (1988).
3. Strobel, S. A. & Dervan, P. B. *Science* **249**, 73-75 (1990).
4. Maher, L. J., Wold, B. & Dervan, P. B. *Science* **245**, 725-730 (1989).
5. Hanvey, J. C., Shimizu, M. & Wells, R. D. *Nucleic Acids Res.* **18**, 157-161 (1990).
6. Barlow, D. P. & Lehrach, H. *Trends Genet.* **3**, 167-171 (1987).
7. Wenzlau, J. M., Saldanha, R. J., Butow, R. A. & Perlman, P. S. *Cell* **56**, 421-430 (1989).
8. Delahodde, A. et al. *Cell* **56**, 431-441 (1989).
9. Patel, Y., Van Cott, E., Wilson, G. G. & McClelland, M. *Nucleic Acids Res.* **18**, 1603-1607 (1990).
10. Weil, M. D. & McClelland, M. *Proc. natn Acad. Sci. U.S.A.* **86**, 51-55 (1989).
11. Koob, M., Grimes, E. & Szybalski, W. *Science* **241**, 1084-1086 (1988).
12. Koob, M. & Szybalski, W. *Science* **250**, 271-273 (1990).
13. Praseuth, D. et al. *Proc. natn Acad. Sci. U.S.A.* **85**, 1349-1353 (1988).
14. Lyamichev, V. I. et al. *Nucleic Acids Res.* **16**, 2165-2178 (1988).
15. Felsenfeld, G., Davies, D. R. & Rich, A. *J. Am. chem. Soc.* **79**, 2023-2024 (1957).
16. Lipsett, M. N. *J. biol. Chem.* **239**, 1256-1260 (1964).
17. Rajagopal, P. & Feigon, J. *Nature* **339**, 637-640 (1989).
18. Povsic, T. J. & Dervan, P. B. *J. Am. chem. Soc.* **111**, 3059-3061 (1989).
19. Griffin, L. C. & Dervan, P. B. *Science* **245**, 967-971 (1989).
20. Horne, D. A. & Dervan, P. B. *J. Am. chem. Soc.* **112**, 2435-2437 (1990).
21. Andreadis, A., Hsu, Y. P., Kohli, G. B. & Schimmel, P. *Cell* **31**, 319-325 (1982).
22. Mortimer, R. K. & Schild, D. *Microbiol. Rev.* **49**, 181-212 (1985).
23. Carle, G. F. & Olson, M. V. *Proc. natn Acad. Sci. U.S.A.* **82**, 3756-3760 (1985).
24. Schwartz, D. C. & Cantor, C. R. *Cell* **37**, 67-75 (1984).
25. Carle, G. F. & Olson, M. V. *Nucleic Acids Res.* **12**, 5647-5664 (1984).
26. Keesey, J. K., Bigelis, R. & Fink, G. R. *J. biol. Chem.* **254**, 7427-7433 (1979).
27. Maher, L. J., Dervan, P. B. & Wold, B. *Biochemistry* **29**, 8820-8826 (1990).

ACKNOWLEDGEMENTS. We thank J. Hanish for helpful discussions. The work was supported by the National Institutes of Health and the Howard Hughes Medical Institute (predoctoral fellowship to S.A.S.).

ERRATUM

Meteoritic silicon carbide: pristine material from carbon stars

Roy S. Lewis, Sachiko Amari & Edward Anders

Nature **348**, 293-302 (1990)

IN this article, the second affiliation for Edward Anders was omitted. In addition to the Chicago address, he is attached to the Physikalisches Institut der Universität, CH-3012 Bern, Switzerland.

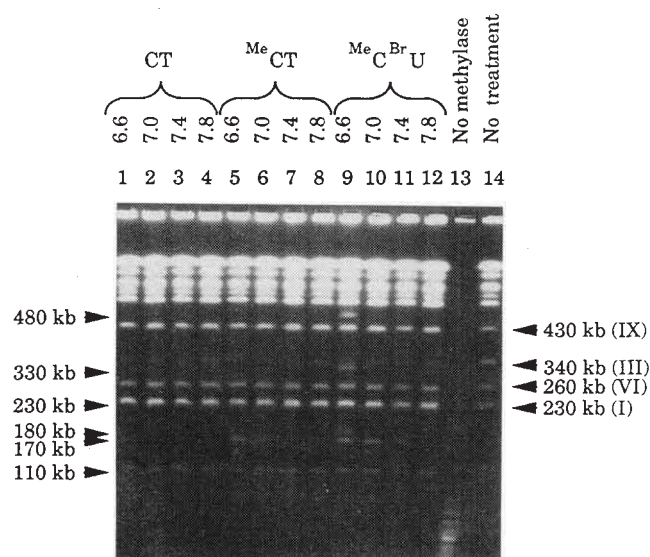


FIG. 4 Triple-helix-mediated enzymatic cleavage of the yeast genome as a function of oligonucleotide composition and pH. Lanes 1-12, reactions on a yeast strain containing the triple helix target site with oligonucleotides (CT, ^{Me}CT and ^{Me}CBrU) and pH values (6.6, 7.0, 7.4 and 7.8) as indicated above figure. Methylation time was 4.5 h. 170-kb and 650-kb (unresolved) secondary cleavage products were observed with both ^{Me}C substituted oligonucleotides at or below pH 7.4. The cleavage site can be assigned to chromosome II (820 kb) by two-dimensional pulsed-field gel electrophoresis in which the DNA was triple helix-protected and methylated before the first dimension and restriction enzyme-digested before the second dimension (data not shown). Additional secondary cleavage sites were observed with the ^{Me}CBrU oligonucleotide at pH 6.6 (lane 9). The 180- and 480-kb products were assigned to chromosome XI (660 kb), and the 330 and 780 kb (unresolved) products were assigned to the VII/XV doublet (1,100 kb) by the same method (data not shown).